



UNIVERSITÀ
DEGLI STUDI
FIRENZE

DOTTORATO DI RICERCA

IN

INFORMATICA, SISTEMI e TELECOMUNICAZIONI

TELEMATICA E SOCIETÀ DELL'INFORMAZIONE

Ciclo XXVII

Coordinatore Prof. CHISCI LUIGI

LA FORMA DELL'ONTOLOGIA DELL'AZIONE IMAGACT: DAL MODELLO
GERARCHICO AL MODELLO INSIEMISTICO IN UN DB A GRAFO

*Rappresentazione dei concetti attraverso prototipi, formati machine readable (OWL) e sfruttamento in task di disambiguazione delle
lingue naturali*

Settore Scientifico Disciplinare ING-INF/05

Dottorando

Dott. GREGORI LORENZO

Tutore

Prof. NESI PAOLO

Prof. MONEGLIA MASSIMO

Coordinatore

Prof. CHISCI LUIGI

ANNI 2012/2014

INDICE

1	INTRODUZIONE	5
2	RISORSE LINGUISTICHE E ONTOLOGIE	12
2.1	Framework generali per la pubblicazione di risorse in RDF	14
2.2	Risorse linguistiche: stato dell'arte	17
2.2.1	WordNet	17
2.2.2	FrameNet, VerbNet e PropBank	20
2.2.3	DOLCE e SUMO	22
2.2.4	PAROLE/SIMPLE	23
2.2.5	BabelNet	24
2.3	Le ontologie in contesto multilingue	25
3	L'ONTOLOGIA DELL'AZIONE IMAGACT	28
3.1	Concetti di base	28
3.1.1	I linguaggi naturali e l'ontologia dell'azione: i verbi generali, l'equivalenza locale e la produttività nelle traduzioni	28
3.1.2	La produttività dei concetti: estensioni primarie e marcate dei verbi: Identificazione e definizione	31
3.2	Il framework di IMAGACT	35
3.2.1	La Strategia corpus-based	35
3.2.2	Il processo induttivo da corpus	38
3.2.3	Il mapping dei tipi su una ontologia interlinguistica dell'azione e la rappresentazione per prototipi	41
3.3	Similarità pragmatica e semantica tra azioni	46
3.4	Le inferenze	52
4	IL MODELLO GERARCHICO DI IMAGACT	54
4.1	Entità e relazioni principali	54

Indice

4.2	Derivazione del modello gerarchico	57
4.3	Proprietà del modello gerarchico	60
4.3.1	La Scena: concetto e rappresentazione	60
4.3.2	Equivalenza locale e famiglie di prototipi	61
4.3.3	Scene vuote	63
4.3.4	Scene figlie di due scene	65
4.3.5	Le inferenze	66
5	LIMITI DEL MODELLO GERARCHICO	69
5.1	Errori di rappresentazione nella struttura dati	69
5.1.1	Caso 1. Costruzione gerarchica ad hoc	70
5.1.2	Caso 2. Mapping mancato tra tipi	73
5.1.3	Caso 3. Difficoltà di modifica per tipi molto generali	75
5.2	Estensione dell'ontologia: casi di studio sulla lingua cinese	77
5.2.1	Caso 1. Mancanza di un verbo generale	78
5.2.2	Caso 2. Riferimento all'azione da diversi punti di vista	81
5.2.3	Caso 3. Azioni senza scopo	82
5.2.4	Caso 4. Azioni senza verbo	83
5.2.5	Caso 5. Movimento fuori dal verbo	84
5.2.6	Caso 6. Mancanza del concetto azionale	85
5.3	Necessità empiriche derivanti dai casi di studio	86
5.3.1	Aggiungere nuovi tipi	86
5.3.2	Aggiungere nuove scene	88
6	RISTRUTTURAZIONE DELL'ONTOLOGIA	90
6.1	Il modello insiemistico	90
6.1.1	La semplificazione dei concetti di IMAGACT	91
6.1.2	Un nuovo modello inferenziale	93
6.2	L'adeguatezza del modello	99
6.2.1	Esempi sui casi di studio	101
7	IMPLEMENTAZIONE IN UN DATABASE A GRAFO	106

Indice

7.1	NoSQL database	107
7.1.1	Tipologie dei database NoSQL	109
7.1.2	Il database a grafo per IMAGACT	111
7.2	Il nuovo DB	113
7.2.1	Neo4J	113
7.3	Questioni di migrazione	117
7.3.1	Rappresentazione dei verbi	118
7.3.2	Rappresentazione delle frasi	119
7.3.3	Le scene vuote	120
7.4	IMAGACT su Neo4J	123
7.4.1	Il database MySQL d'origine	123
7.4.2	Implementazione della struttura dati in Neo4J	127
7.4.3	Migrazione del DB	131
8	VALIDAZIONE DELL'ONTOLOGIA	139
8.1	Il quadro della validazione di IMAGACT	139
8.1.1	Validazioni rispetto ai sistemi di traduzione automatica e disambiguazione	144
8.1.2	La validazione dei dati in input e output dell'ontologia IMAGACT	146
8.1.2.1	Test di identificazione della variazione primaria e marcata	148
8.1.2.2	Test di riconoscimento dei tipi azionali	151
8.2	I concetti azionali in BabelNet e IMAGACT	155
8.3	Costruzione del dataset di riferimento per la validazione di IMAGACT	160
8.4	Valutazione della copertura dei traduttori	163
8.5	La traduzione dei concetti azionali in BabelNet e Google Translate	165
8.6	Relazioni tra i risultati dei test precedenti	170
9	VERSO UN'ONTOLOGIA WEB	177
9.1	Costituzione dell'ontologia IMAGACT-OWL	177
9.2	Mapping su <i>lemon</i>	180

Indice

9.3	Collegamento IMAGACT-BabelNet	183
9.3.1	Tipi azionali e BabelSynset	185
9.3.2	Test di <i>linking</i> tra le ontologie	185
	Bibliografia	193

Appendici

disponibili all'indirizzo <http://lablita.dit.unifi.it/app/~gregori/phd/>

INTRODUZIONE

Un verbo di azione può riferirsi, nel suo significato proprio, a molte azioni diverse e non possiamo aspettarci che le stesse alternative siano possibili per il verbo che lo traduce in un'altra lingua. Corrispondentemente i sistemi di traduzione automatica hanno problemi specifici proprio quando debbono tradurre frasi semplici costruite con verbi d'azione. Per esempio se si chiede al traduttore di Google di tradurre in italiano la frase inglese seguente

“John pushes the garbage into a can”

otteniamo “John spinge l'immondizia nel secchio”

invece di “John comprime l'immondizia nel secchio”.

Andando nell'altra direzione, ovvero dall'italiano all'inglese,

“Mario prende il gatto per la coda”

è tradotto con “Mario takes the cat by the tail”

invece di “Mario catches the cat by the tail”, che corrisponde alla sua reale interpretazione.

Allo stesso modo:

“Mario gira gli zucchini nella padella”

è tradotto con “John turns the zucchini in the pan”

invece di “John stirs the zucchini in the pan”.

La ragione principale di questi errori è che la serie delle interpretazioni possibili per i verbi di azione *prendere*, *spingere*, *girare*, *comprimere* in italiano e *take*, *catch*, *push* e *turn* in inglese, e le loro corrispondenze con le azioni vere e proprie, non sono presenti nel sistema e in realtà non sono a disposizione in modo chiaro né nei dizionari né nelle

moderne ontologie linguistiche. Non è quindi possibile disambiguare il verbo per fornire la traduzione adatta all'azione riferita nella frase originale.

Il capitolo 2 della tesi presenta lo stato dell'arte delle principali ontologie linguistiche, quali WordNet e BabelNet, che sono risorse correntemente sfruttate per la disambiguazione. Saranno presentati inoltre i quadri generali di rappresentazione dell'informazione ontologica e per la pubblicazione di risorse in RDF del dato linguistico, ad esempio *lemon* e SKOS.

Il problema della disambiguazione dei verbi di azione è significativo, perché il riferimento alle azioni è molto frequente nell'uso linguistico di ogni giorno e proprio i verbi che si usano più spesso si riferiscono, in effetti, a molti tipi di azioni diverse e sono chiamati per questo, come vedremo, "generali".

Il progetto IMAGACT, realizzato nel programma PAR/FAS 2007-2013 della regione Toscana, ha prodotto una infrastruttura linguistica basata su corpora che identifica le azioni più frequentemente riferite nel parlare quotidiano. Analizzando specificamente grandi corpora orali di lingua italiana e inglese sono state identificate 1.010 azioni diverse riferite dai parlanti nelle loro interazioni quotidiane e tali azioni sono state rappresentate con scene prototipiche, filmate in studio o realizzate come animazioni 3D. Le corrispondenze di queste azioni con i verbi che possono riferirsi ad esse in inglese, italiano, spagnolo e cinese sono state definite in un database accessibile on-line all'indirizzo www.imagact.it

Il capitolo 3 di questa tesi dà conto in breve dei concetti linguistici di base necessari al trattamento semantico dei verbi di azione e delle modalità con cui IMAGACT è stata sviluppata. In sintesi, a valle del lavoro di sviluppo, IMAGACT dà un quadro della varietà di attività che sono prominenti nella vita quotidiana e per ognuna di esse specifica il verbo o i verbi che la esprimono in lingue ad alto impatto a livello globale (italiano, inglese, spagnolo, cinese). IMAGACT rende conto di due proprietà centrali della rappresentazione dell'azione nelle lingue naturali: ogni verbo può esprimere uno o più concetti azionali, mentre ogni concetto può essere identificato da uno o più verbi. Le intersezioni e gli insiemi complemento dei riferimenti alle azioni dei verbi sono evidenziate sia a livello intralinguistico che e interlinguistico.

IMAGACT contiene varie innovazioni rilevanti nel quadro delle ontologie linguistiche. La prima di queste è la rappresentazione delle entrate, ovvero dei concetti di azione individuati, attraverso scene che rappresentano un prototipo dell'azione; la seconda innovazione è la strutturazione delle relazioni di inferenza concettuale all'interno dell'ontologia, che sono basate sulla relazione pragmatica di *equivalenza locale*.

La rappresentazione dei concetti attraverso scene prototipali comporta che questi possono essere facilmente identificati dagli utenti umani indipendentemente dalla loro lingua e quindi permette una facile estensione a nuove lingue.

Per esempio le scene corrispondenti ai seguenti *thumbnail* rappresentano la variazione del verbo *girare* su più tipi di azione diversi che vanno dal "movimento nello spazio", al "movimento di un soggetto su se stesso", a più tipi di azione sull'oggetto: "rotazione", "capovolgimento", "orientamento".

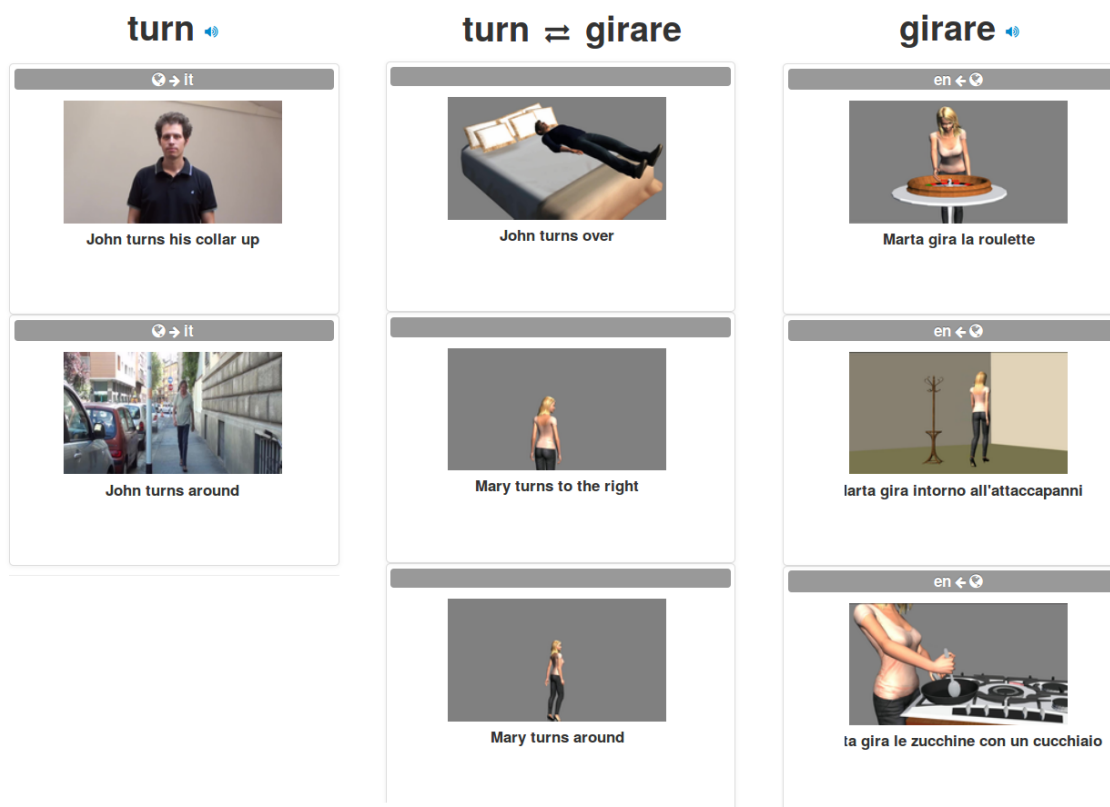
Figura 1: La variazione di *girare* per tipi di azione



Un verbo come *girare* si dice appunto "generale" in quanto si riferisce ad azioni tra loro pragmaticamente diverse. Crucialmente in ogni Tipo azionale cambiano le equivalenze linguistiche del verbo, e tali variazioni costituiscono differenziali linguistici per ciascun Tipo. Per esempio nella figura evidenziata in rosso il verbo è "equivalente" a *curvare* (ovvero lo implica), in quella evidenziata in blu il verbo *girare* è equivalente a *rovesciare*, mentre in quella in verde a *mescolare*. Tali equivalenze valgono solo localmente, ma sono sistematiche e vanno a definire i rispettivi concetti azionali, determinando le loro relazioni inferenziali. I verbi delle lingue non coprono nello stesso modo l'universo delle azioni, ovvero risultano traducibili vicendevolmente solo su una parte delle azioni a cui si estendono. Per esempio *girare* e il corrispondente inglese *turn* covariano in gran

parte dei tipi di azione in questione, ma sistematicamente non covariano quando *girare* è equivalente a *mescolare* e non quando *turn* è equivalente a *rigirarsi* o *rivoltarsi*. La figura 2 visualizza questo differenziale.

Figura 2: Comparazione di *turn* vs. *girare*



La possibilità di stabilire differenziali in estensione dei verbi generali costituisce una pietra miliare sia per permettere la disambiguazione della frasi e la loro traduzione sia per la modellazione dei concetti azionali, anche perché le moderne ontologie linguistiche quali WordNet o BabelNet sono fortemente deficitarie nello specificare tali variazioni e quindi non consentono a priori alcuna disambiguazione nel dominio dell'azione.

Di tale insufficienza si dà conto nel capitolo 8 in cui, dopo aver illustrato le varie prove di validazione condotte per garantire la qualità dell'ontologia, si valida la rilevanza di IMAGACT specificamente come risorsa per la disambiguazione e la traduzione. In particolare, con un esperimento di validazione condotto in questa tesi sulla base di un data set statisticamente significativo, mostreremo in modo obiettivo che le variazioni semantiche dei verbi d'azione più frequenti nell'uso linguistico non sono discriminate

in BabelNet e inoltre non sono tradotte correttamente in Google Translate, due risorse che sono a tutt'oggi l'ontologia più ricca a disposizione e il sistema più popolare di traduzione automatica. Sebbene l'identificazione di una serie ampia di variazioni possibili di interpretazione pragmatica dei verbi di azione sia un passo necessario per permettere la disambiguazione delle frasi naturali, ciò non è sufficiente, sia perché i concetti rappresentati iconicamente non sono leggibili da una macchina, sia perché la quantità di informazione linguistica associata ad ogni concetto è scarsamente significativa a livello statistico. Di questi aspetti e delle strategie per l'implementazione dei dati disponibili in formati *machine readable* si dà conto nel capitolo 9.

Date queste premesse, la tesi si concentra però sulla forma dell'ontologia IMAGACT e sulle difficoltà che il modello classico delle ontologie gerarchiche, che costituisce la forma attuale anche di questa ontologia, incontra quando si debbano trattare sia un gran numero di variazioni semantiche dei verbi di molte lingue diverse sia le relazioni inferenziali tra concetti rappresentati attraverso prototipi. Tali relazioni, per essere empiricamente adeguate al trattamento delle lingue naturali, non debbono valere in tutte le applicazioni di un verbo, ovvero non sono globali, ma valgono solo localmente e devono rendere conto delle relazioni di equivalenza locale del verbo in ciascuna delle sue interpretazioni. Per esempio le inferenze tra *girare* (verbo più generale) e *curvare* (verbo più specifico) e quelle tra *girare* (verbo più generale) e *rovesciare* / *capovolgere* / *voltare* / *rivoltare* (verbi più specifici) valgono ciascuna in una soltanto delle interpretazioni del verbo, ma nondimeno sono essenziali per la caratterizzazione delle relazioni tra concetti nella sfera dell'azione.

Figura 3: Esempio di equivalenze locali in *girare*



Ciò comporta però che le rappresentazioni prototipali che identificano, ad esempio le due entrate in considerazione siano adeguate a rappresentare l'azione dal punto di vista di tutti i verbi equivalenti e a consentire quindi una estensione produttiva adeguata di entrambi i concetti verbali sulle istanze possibili del prototipo. La soluzione dovrà

inoltre essere adeguata ad esprimere le equivalenze nei concetti azionali non solo del lessico italiano, ma anche del lessico delle altre lingue; ad esempio per l'inglese, oltre a *turn*, nel primo caso è equivalente il verbo *round* e nel secondo il verbo *flip*.

Questo requisito comporta quindi una complessità di problemi teorici relativi alla rappresentazione dei concetti attraverso prototipi e necessita soluzioni formali e tecniche che costituiscono il centro di questa tesi di dottorato. La nozione di prototipo non si dimostra infatti sufficiente a specificare le relazioni inferenziali che si realizzano localmente. In particolare il concetto di "famiglia di prototipi" risulta essere più adeguato a cogliere la complessità dei concetti linguistici e le loro relazioni. Per esempio nel caso di "girare la cartolina" possiamo pensare che anche "girare una tazza" cada nell'estensione dello stesso concetto, ma a questa azione, che può costituire il prototipo del verbo *rovesciare*, non è applicabile *voltare*. In altri termini: dal punto di vista di *voltare* i due prototipi costituiscono un differenziale, mentre dal punto di vista della variazione di *rovesciare* i due prototipi costituiscono due tipi diversi di rovesciamento (uno che implica *voltare* e uno no); infine dal punto di vista del verbo generale *girare* ricadono entrambi in un solo concetto, la cui granularità deve essere minore, perché serve a raggruppare le rotazioni sull'asse orizzontale, stabilendo un differenziale con i movimenti nello spazio e le rotazioni sull'asse verticale.

Figura 4: Esempio di equivalenze locali in *girare*



L'entrata nell'ontologia, in relazione al verbo *girare*, non potrà essere quindi in questo caso rappresentata da un prototipo, ma piuttosto da una collezione di prototipi (famiglia), nessuno dei quali risulta più emblematico dell'altro a caratterizzare il concetto, ma che incorporano i differenziali linguistici che abbiamo descritto e una gerarchia implicita, dal più generale al più specifico.

Nel capitolo 4 si illustra la forma gerarchica dell'ontologia IMAGACT così come è stata costruita e proiettata in un DB MySQL. In estrema sintesi dimostreremo che tale forma porta da un lato a soluzioni fortemente *ad hoc* e dall'altro a una strutturazione rigida che non risponde alle necessità di implementazione e modificazione rispetto a nuovi concetti e a diversi modi di segmentare l'esperienza che l'ontologia deve incorporare quando nuove lingue sono sviluppate in essa. Di tali rigidità nella strutturazione dell'ontologia si dà conto nel capitolo 5 portando un'ampia serie di casi sia di carattere formale sia derivati dall'integrazione nella risorsa di lingue diverse dall'italiano e dall'inglese, e in particolare della lingua cinese. Nel capitolo 6 proporremo una diversa modellazione delle relazioni nell'ontologia e vedremo in particolare come la sostituzione del modello gerarchico con un modello insiemistico porti a risolvere i principali problemi di rappresentazione e implementazione.

Nel capitolo 7 verrà descritta una re-implementazione dell'ontologia in un DB a grafo, soluzione questa che, rispetto a un DB MySQL, si presta da un lato a una più diretta relazione con il modello insiemistico e dall'altro a una conversione più diretta in un formato *machine readable* (OWL). Nel capitolo saranno illustrati in particolare i passi tecnici compiuti per migrare il DB MySQL in un DB Neo4J. Nel capitolo 9 infine sarà trattata la questione dello sfruttamento dell'ontologia in ambito web: in particolare verrà proposta l'esportazione del nucleo del DB a grafo in un modello RDF-OWL e la valutazione di un mapping su BabelNet. A questo proposito vedremo come le differenze strutturali tra le due ontologie rendano difficilmente percorribile la strada di un mapping tra entità, ma come invece l'approccio di *instance matching* risulti particolarmente fruttuoso per il collegamento delle due risorse.

RISORSE LINGUISTICHE E ONTOLOGIE

Le risorse linguistiche in rete possono essere classificate in due macro-categorie: risorse semantico-lessicali e corpora.

I corpora sono collezioni di materiale testuale raccolti sulla base di specifici criteri costitutivi (*corpus design*) per essere rappresentativi di un dominio linguistico. Essi costituiscono la base per condurre studi empirici sull'uso della lingua; inoltre sono anche sfruttabili nell'ambito del trattamento automatico del linguaggio, in particolare per addestrare algoritmi di riconoscimento di fenomeni linguistici.

Un corpus è una raccolta strutturata di eventi comunicativi prodotti in ambiente naturale e selezionati sulla base di criteri espliciti al fine di rappresentare una lingua o una sua specifica varietà.

[Cresti and Panunzi, 2013]

I corpora possono essere molto diversi tra loro per vari parametri come la lingua, le dimensioni, il dominio, lo scopo per il quale sono stati realizzati; inoltre possono essere di lingua scritta o di lingua parlata. L'importanza dei corpora per la ricerca linguistica è testimoniata dalla grande tradizione di studi, la *corpus linguistics*, che risale ad oltre un secolo fa, ma che si è sviluppata in modo consistente a partire dagli anni '50, con la costruzione di corpora elettronici (Busa, 1951) e l'impiego di calcolatori, in grado di analizzare grandi quantità di materiale testuale in poco tempo.

Le risorse lessicali sono un'ampia classe di oggetti nei quali l'informazione linguistica è specificata a livello di parola; di questa categoria fanno parte i dizionari, le tassonomie, i thesauri e le ontologie, ciascuno con un diverso grado di strutturazione informativa. Ad eccezione dei dizionari, che non contengono relazioni semantiche tra lemmi, le altre risorse vengono dette semantico-lessicali, poiché collegano le entrate in una rete semantica, che può essere una gerarchia, nel caso delle tassonomie, oppure un grafo, nel caso di thesauri e ontologie. Sebbene la differenza tra queste ultime due non sia

sempre così chiara, possiamo pensare ad un thesaurus come una rete di parole e ad una ontologia come una rete di concetti: naturalmente entrambe le risorse contengono sia parole che concetti, ma, mentre il thesaurus ha alla base i lemmi e quindi può essere considerato come un dizionario arricchito di relazioni semantiche, l'ontologia è una concettualizzazione del mondo, che quindi si fonda su concetti collegati in una rete di relazioni, ai quali poi possono essere legate parole.

Come definizione di ontologia riportiamo la seguente, una delle più utilizzate in letteratura:

An ontology is a formal, explicit specification of a shared conceptualization
[Studer et al., 1998]

Corpora e risorse lessicali non sono due mondi completamente separati e indipendenti; infatti normalmente le risorse lessicali sono realizzate a partire dalle informazioni estratte da corpus e parallelamente i corpora si collegano alle risorse lessicali quando vengono arricchiti da informazioni semantiche a livello di parola. A titolo di esempio citiamo il corpus MASC (Ide et al., 2010), un corpus di inglese americano di 500mila token derivato dal Open American National Corpus¹, arricchito con l'annotazione manuale dei sensi di parola di WordNet, e IMAGACT, l'ontologia dell'azione derivata dall'annotazione di corpora di parlato che costituisce l'oggetto di questa tesi.

Le risorse linguistiche trovano un'importante area applicativa nell'elaborazione automatica del linguaggio naturale (*Natural Language Processing*, NLP): far comprendere ad una macchina il linguaggio umano è infatti un obiettivo di ricerca ambizioso, ma fortemente perseguito in questi anni in cui la tecnologia è sempre più pervasiva della vita quotidiana delle persone e la domanda di interazione naturale uomo-macchina è crescente. Senza dilungarci sulla schiera sterminata di ambiti applicativi in cui il trattamento automatico del linguaggio trova sbocco, diciamo che quello di elaborare il linguaggio è un task molto complesso, proprio a causa di alcune caratteristiche dei linguaggi naturali, come vaghezza e ambiguità (Wasow et al., 2005; Martin, 2007), che li rendono difficili da trattare in modo algoritmico. In particolare alcune azioni che una persona può facilmente eseguire grazie alla sua competenza linguistica - come sintetizzare il contenuto di un

¹ <http://www.anc.org/>

testo (*Summarization*), fare una traduzione (*Machine Translation*), identificare il significato corretto di una parola in un contesto (*Word Sense Disambiguation*) o partecipare a una conversazione (*Dialogue Systems*) - sono estremamente ardue per una macchina. Tutti questi compiti richiedono infatti l'interpretazione del linguaggio e un'interpretazione necessita della conoscenza del mondo (Ovchinnikova, 2012); pertanto affiancare agli algoritmi ontologie e basi di conoscenza (*Knowledge Base, KB*) è la strada che viene correntemente utilizzata per affrontare questi task complessi in modo automatico, che altrimenti risulterebbero intrattabili.

Lo sviluppo di tecnologie web ha portato alla definizione di linguaggi condivisi per la rappresentazione della conoscenza in rete; in particolare ad oggi RDF e suoi derivati (RDFS, OWL) costituiscono lo standard formale per la pubblicazione web di ontologie e basi di conoscenza. Pubblicare le risorse linguistiche in RDF significa da un lato avere un formato standard di rappresentazione dei dati, dall'altro rendere disponibile la risorsa in rete come Open Data per lo sfruttamento in task avanzati di NLP. In questo ambito è particolarmente rilevante l'iniziativa del *Linguistic Linked Open Data Cloud*² (Chiarcos et al., 2011), che raccoglie e collega ontologie linguistiche in RDF e ad oggi contiene più di 500 risorse (Fig. 5).

2.1 FRAMEWORK GENERALI PER LA PUBBLICAZIONE DI RISORSE IN RDF

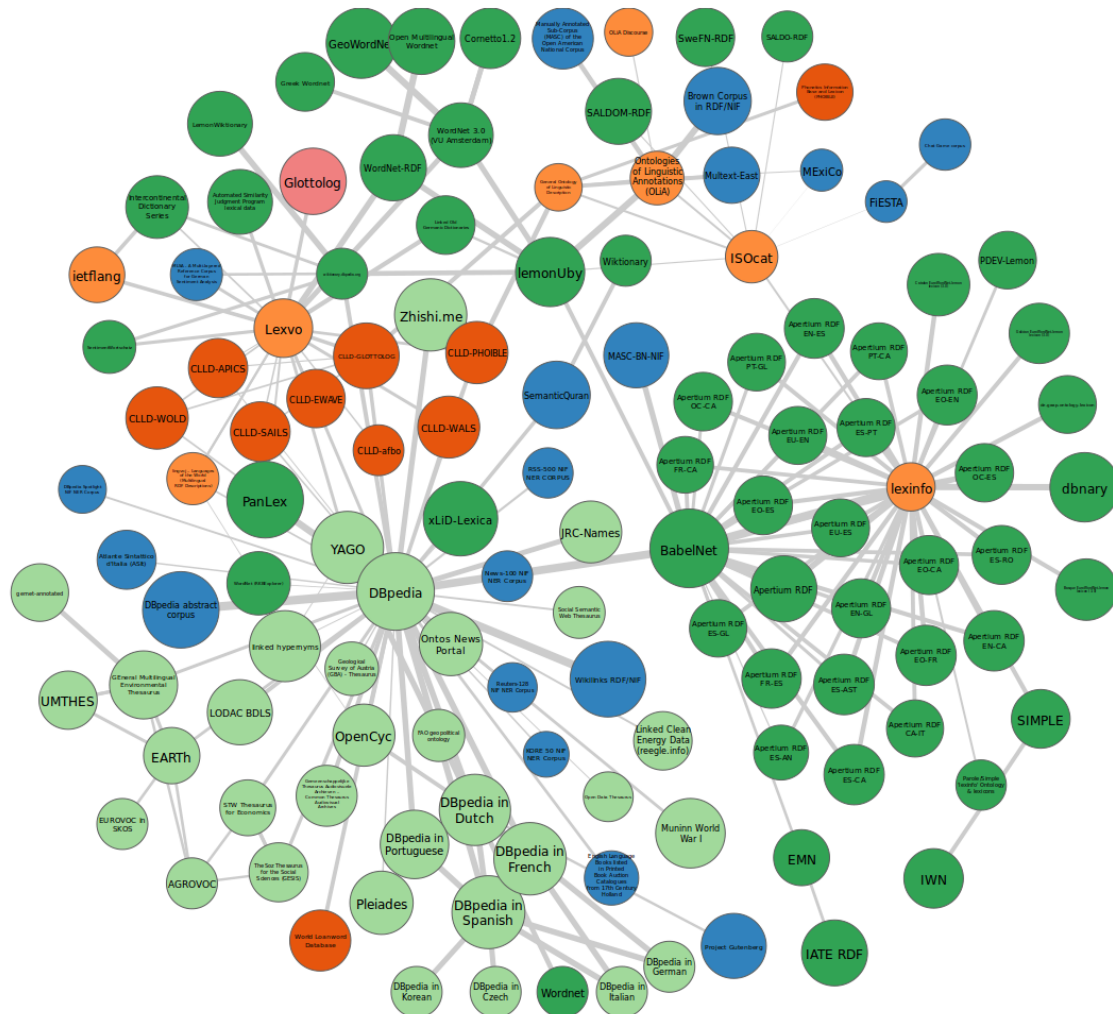
In questo paragrafo descriviamo brevemente due formalismi particolarmente interessanti per la pubblicazione di risorse lessicali in rete: SKOS e *lemon*.

SKOS (Simple Knowledge Organization System)³ è un modello dati che consente una rappresentazione standard in RDF di risorse lessicali quali thesauri, tassonomie e dizionari: in particolare tutte quelle risorse che non sono in forma di ontologia possono essere portate in RDF utilizzando i formalismi di SKOS. Questo framework, nato nel 2003 e mantenuto dal W3C, mette a disposizione una serie di concetti e relazioni di base in RDF e RDFS, all'interno dello SKOS Core Vocabulary (Miles and Brickley, 2005). Il modello di SKOS è molto semplice, contiene poche definizioni di alto livello: le due classi principali

² <http://linguistic-lod.org/llod-cloud>

³ <http://www.w3.org/2004/02/skos/>

Figura 5: Grafo delle risorse linguistiche in RDF disponibili nel Linguistic Linked Open Data Cloud.



sono *Concept*, unità di base all'interno del quale rientrano gli oggetti del thesaurus (parole, significati, categorie,...) e *ConceptScheme*, che racchiude più *Concept*. Ci sono poi tre relazioni semantiche principali che possono collegare i *Concept* tra loro: *broader*, *narrower* e *related*. Le prime due servono per rappresentare una gerarchia tra concetti, ciascuno dei quali può essere più generale (*broader*) o più specifico (*narrower*) di un altro; *related* invece determina una generica affinità semantica. Avere un modello così semplice ha dei vantaggi in termini di facilità di utilizzo e applicabilità a risorse di diverse tipologie; inoltre, essendo in RDF, il modello può essere esteso collegandolo ad altre ontologie e creando le proprie specializzazioni. D'altra parte una serie di informazioni linguistiche

più fini che sono spesso contenute nei thesauri (come ad esempio la morfologia), non trova una rappresentazione nel modello di SKOS (McCrae et al., 2011).

lemon (Lexicon Model for Ontologies)⁴ è un'ontologia fondazionale (o *top-level*) per la rappresentazione di risorse lessicali in RDF. L'ontologia è il risultato di un progetto che ha coinvolto diverse università europee⁵ ed ha portato alla formalizzazione delle entità e delle relazioni che sono alla base dei lessici. *lemon* è completamente indipendente dalla lingua e può essere sfruttato per collegare e pubblicare risorse linguistiche esistenti, attraverso il mapping sulle entità del *lemon core*, che è il nucleo dell'ontologia (Fig. 6). Un *Lexicon* rappresenta il lessico di una lingua e contiene tutte le entrate lessicali, ciascuna delle quali è un'istanza di *LexicalEntry*. Ogni *LexicalEntry* ha una forma canonica (il lemma) e può avere altre forme di parola (*LexicalForm*) dipendenti da variazioni di genere, numero, tempi e modi verbali, ecc. Inoltre una *LexicalEntry* può avere uno o più *LexicalSense*, che sono propriamente i sensi di parola; questa entità funge da collegamento delle *LexicalEntry* con le entità ontologiche (*Ontology*). Infine non è detto che ogni *LexicalEntry* debba essere un'unica parola, ma può essere un'unità più grande o più piccola: *lemon* mette a disposizione 3 classi descrittive delle *LexicalEntry*, che possono essere parole (*Word*), sequenze di parole (*Phrase*) o porzioni di parola (*Part*).

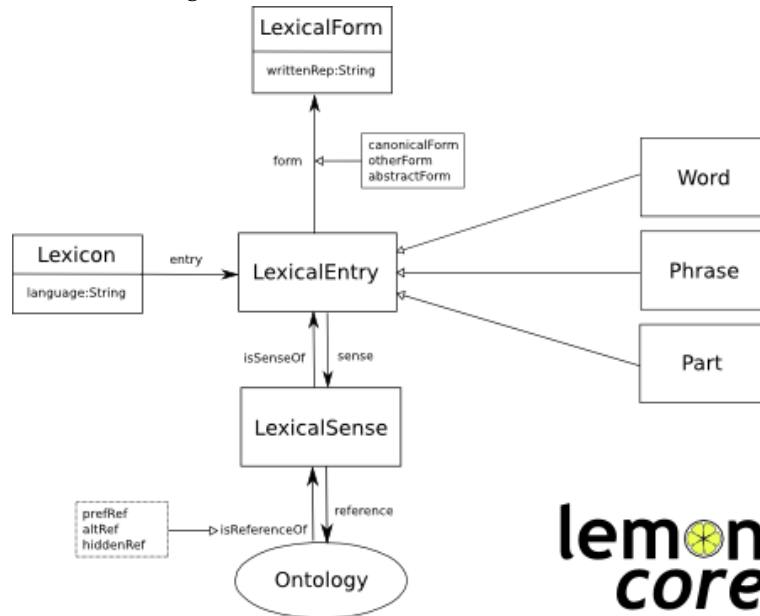
Sopra il *core*, che contiene gli elementi di base, *lemon* fornisce una serie di moduli, per aggiungere altri livelli di informazione linguistica secondo le proprie esigenze di rappresentazione:

- il *Linguistic Description Module* fornisce descrizioni aggiuntive di proprietà di *LexicalEntry* e *LexicalForm*;
- il *Phrase Structure Module* permette di dettagliare la struttura delle sequenze di parole (*Phrase*);
- il *Syntax and Mapping Module* consente di rappresentare gli usi sintattici delle *LexicalEntry* e, di legare i costituenti (ad esempio i frame semantici) a entità ontologiche;

⁴ <http://lemon-model.net>

⁵ Bielefeld University, National University of Ireland, Universidad Politécnica de Madrid e Deutsche Forschungszentrum für Künstliche Intelligenz

Figura 6: La struttura del *lemon core*



- il *Variation Module* descrive relazioni sussistenti tra diverse entrate (ad esempio una relazione di traduzione);
- il *Morphology Module* permette di aggiungere informazione morfologica.

Grazie alla struttura modulare quindi *lemon* consente una rappresentazione più dettagliata delle strutture linguistiche rispetto a SKOS e si sta affermando come standard per la pubblicazione di risorse lessicali sul web.

2.2 RISORSE LINGUISTICHE: STATO DELL'ARTE

2.2.1 *WordNet*

La risorsa semantico-lessicale che ha avuto il maggior impatto nell'ambito della disambiguazione – e in generale in task di elaborazione automatica del linguaggio - è sicuramente WordNet (Miller et al., 1990; Fellbaum, 1998), sviluppato dall'università di Princeton. WordNet è un thesaurus che organizza il lessico inglese in una rete semantica formata da raggruppamenti di parole con lo stesso significato, i *synset*, e da relazioni semantiche tra questi. Ciascun lemma in WordNet (*word form*) può avere uno o più sensi (*word meaning*),

ognuno dei quali è identificato da un numero; *word meaning* diversi che fanno riferimento a significati affini sono raggruppati all'interno dello stesso *synset* che rappresenta propriamente un concetto, il quale è esplicitato sia da una glossa, sia da tutte le parole, o meglio i *word meaning*, in esso contenuti. L'insieme dei *synset* costituisce quindi l'inventario dei concetti presenti in inglese con tutte le sue possibili lessicalizzazioni. Inoltre i *synset* in WordNet non sono isolati, ma collegati tra loro attraverso relazioni semantiche (iperonimia, iponimia, meronimia, ecc.) che permettono di esprimere i rapporti tra concetti. In WordNet sono presenti 27 relazioni semantiche distribuite nelle 4 categorie sintattiche trattate: nomi, verbi, aggettivi e avverbi (Tab. 2.2.1).

Uno strumento di *Word Sense Disambiguation* è un algoritmo che assegna un senso ad una parola a partire da una sua occorrenza, ossia dalle informazioni ricavabili dal contesto in cui è applicata. Per fare questo è necessario avere l'inventario dei sensi possibili di una parola (*Word Sense Discrimination*) e una funzione di probabilità che assegni un peso ad ogni senso sulla base del contesto. Una risorsa come WordNet è quindi fondamentale per questo task, poiché contiene sia la lista dei sensi possibili di ogni lemma, sia il contesto semantico di ogni concetto, determinato dalla rete di relazioni con gli altri. Un algoritmo può quindi calcolare la probabilità di ogni *synset* possibile (che è un insieme discreto e finito di candidati) basandosi sul confronto tra le *features* dell'occorrenza e le relazioni semantiche di ciascun *synset*. Avere un buon algoritmo di disambiguazione è fortemente dipendente da correttezza e completezza dell'inventario dei sensi di parola e delle relazioni che li legano. Come vedremo (Cap. 8) il caso dei verbi generali d'azione è problematico da questo punto di vista perché i sensi di parola non sono ben discriminati in risorse come WordNet (e BabelNet).

A partire da WordNet sono nate molte altre risorse e progetti che, a vari livelli, sono delle estensioni alla risorsa di Princeton:

- estensioni ad altre lingue: di WordNet ne esistono diverse versioni specifiche per lingue diverse dall'inglese; per l'italiano c'è ad esempio ItalWordnet (Roventini et al., 2000), sviluppato dall'Istituto di Linguistica Computazionale del CNR di Pisa;
- estensioni a domini specifici: per arricchire WordNet con thesauri più specializzati

Tabella 1: Le relazioni semantiche in WordNet

Nomi	Verbi	Aggettivi	Avverbi
!	! Antonym	!	! Antonym
@	@ Hypernym	&	\ Derived from adjective
@i	~ Instance Hypernym	<	;c Domain of synset - TOPIC
~	* Hyponym	\	;r Domain of synset - REGION
~i	> Instance Hyponym	=	;u Domain of synset - USAGE
#m	^ Member holonym	^	
#s	\$ Substance holonym	;c	
#p	;c Part holonym	;r	
%m	;r Member meronym	;u	
%s	;u Substance meronym		
%p	! Part meronym		
=	= Attribute		
+	+ Derivationally related form		
;c	;c Domain of synset - TOPIC		
-c	-c Member of this domain - TOPIC		
;r	;r Domain of synset - REGION		
-r	-r Member of this domain - REGION		
;u	;u Domain of synset - USAGE		
-u	-u Member of this domain - USAGE		

sono nate molte versioni di WordNet di dominio, che si mappano sulla struttura centrale attraverso speciali relazioni *plug-in*. Jur-WordNet (Sagri et al., 2004) e ArchiWordNet (Bentivogli et al., 2004) sono due esempi di WordNet italiani, che focalizzano rispettivamente sul dominio giuridico e su quello dell'architettura e delle costruzioni;

- estensioni multilingui: MultiWordnet, EuroWordNet e GlobalWordnet⁶ sono iniziative internazionali per la costituzione di una rete multilingue ispirata ai principi di WordNet;
- corpora annotati: l'annotazione di corpora con i *synset* di WordNet risulta particolarmente utile per il training di algoritmi di disambiguazione basati su *machine learning*; esempi sono SemCor (Mihalcea, 1998) o MASC;
- WordNet RDF (McCrae et al., 2014): la versione RDF di WordNet permette la sua integrazione con altre ontologie disponibili in rete;
- *linking* con altre risorse: esistono numerosi tentativi di collegamento di WordNet con altre risorse linguistiche; un esempio su tutti è BabelNet, di cui parleremo più avanti.

2.2.2 FrameNet, VerbNet e PropBank

FrameNet è un'altra risorsa semantico-lessicale in lingua inglese; è stata prodotta dal progetto FrameNet (Baker et al., 1998) e si fonda sulla teoria della Frame Semantics (Fillmore and Atkins, 1992). In tale modello, un'unità lessicale è rappresentata da una coppia "parola-significato". Tipicamente, infatti, ogni senso di una parola polisemica appartiene a Frame semantici differenti. Un Frame è definito da uno script concettuale che descrive un Tipo astratto di situazione o evento, insieme ai suoi partecipanti e ai loro rispettivi ruoli. La risorsa è stata inoltre arricchita attraverso l'annotazione su corpora di lingua inglese. Il database complessivo contiene più di 13.000 unità lessicali (di cui 5.000 verbi) in oltre 1.200 frame semantici.

⁶ Global Wordnet Association (GWA), <http://globalwordnet.org>

VerbNet e PropBank sono due risorse che sfruttano la classificazione di verbi di Levin (1993), in cui si assume che le strutture sintattiche in cui un verbo può occorrere siano conseguenza diretta delle proprietà semantiche soggiacenti e siano pertanto utilizzabili per una classificazione semantica del lessico verbale. Le classi di verbi sono quindi definite a partire da modelli di alternanza sintattica (es. alternanze di diatesi, transitivo/intransitivo, causativo/incoativo): di conseguenza, i membri di ciascuna classe condivideranno sia tratti semantici che proprietà distribuzionali. In VerbNet (Kipper Schuler et al., 2000; Kipper Schuler, 2005) sono estese le classi di Levin con una rappresentazione astratta della struttura sintattica, esplicitando il ruolo semantico dei suoi costituenti. La base di dati consiste in un ampio lessico verbale gerarchizzato (più di 5.000 sensi, corrispondenti a circa 3.700 lemmi verbali divisi in 274 classi), attualmente disponibile solo per l'inglese. Ciascuna classe è costituita da un insieme di verbi che condividono strutture sintattiche, ruoli tematici e restrizioni di selezione. Ogni verbo appartenente a una classe è considerato semanticamente non ambiguo ed è collegato al *synset* di WordNet che rappresenta il suo senso. Inoltre, per descrivere meglio il comportamento semantico di una classe sono stati aggiunti dei *semantic predicates*, che catturano relazioni di diverso Tipo, come quelle espresse in WordNet (antonimia e implicazione) o in FrameNet.

Il progetto PropBank (Babko-Malaya et al., 2004; Palmer et al., 2005) ha portato all'annotazione semantica della Penn Treebank a diversi livelli: struttura argomentale dei verbi, ruoli semantici e alternanze sintattico/semantiche. Nella realizzazione del progetto sono state annotate complessivamente più di 110.000 frasi; il lessico risultante contiene circa 3.600 verbi, con dati quantitativi rispetto ai modelli di alternanza che sono stati riscontrati.

SemLink⁷ è un progetto per il collegamento di risorse linguistiche attraverso il mapping tra le entità che le costituiscono. Attualmente le risorse in oggetto sono WordNet, VerbNet, FrameNet e Propbank e in particolare ad oggi (ver. 1.2.2c) sono stati realizzati due mapping: VerbNet-PropBank e VerbNet-FrameNet.

⁷ <https://verbs.colorado.edu/semLink/>

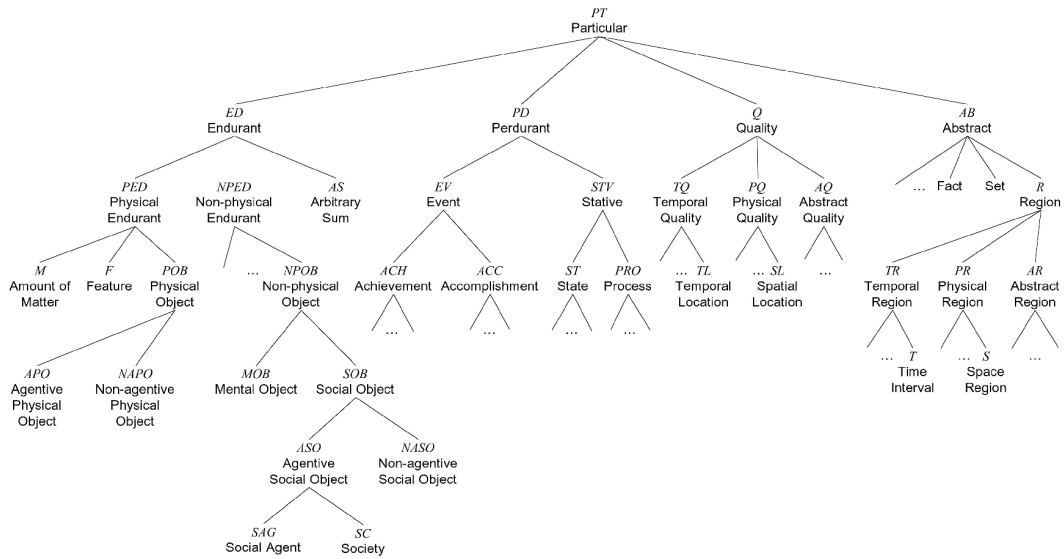
2.2.3 DOLCE e SUMO

Sempre nell'ambito delle *top-ontologies* citiamo due risorse che possono essere considerate storiche: DOLCE e SUMO.

DOLCE (Descriptive Ontology for Linguistic and Cognitive Engineering) è un'ontologia fondazionale uscita nel 2004 dal progetto europeo WonderWeb, che contiene una concettualizzazione generale del mondo derivata dalle scienze cognitive. DOLCE è un'ontologia di particolari (e non di universali) in cui lo scopo è quello di rappresentare gli oggetti del mondo non per la loro natura essenziale, ma così come sono mediati dall'uomo attraverso processi cognitivi di categorizzazione, percezione e interpretazione, ispirandosi alla nozione di *deep background* di Searle (Searle, 1983). La distinzione fondamentale di alto livello in DOLCE è quella tra oggetti *endurant* e *perdurant*. Gli *endurant* sono interamente presenti in ogni istante e possono subire modificazioni nel tempo (oggetti), mentre i *perdurant* sono solo parzialmente presenti in ogni istante e immutabili (eventi). Oltre a queste due categorie, che rappresentano genericamente oggetti ed eventi, ci sono altre due entità di base: *quality* e *abstract*. Le qualità sono le entità che misuriamo e percepiamo (forma, colore, dimensione,...), non sono universali (in questo quadro l'universalità è una caratteristica delle proprietà), ma sono strettamente legate agli oggetti a cui si riferiscono: se l'oggetto smette di esistere, anche la qualità scompare; inoltre due oggetti diversi non possono avere la stessa qualità, ma due qualità diverse con lo stesso valore. Infine gli *abstract* (entità astratte) sono entità che categorizzano gli spazi concettuali. La versione completa di DOLCE contiene 77 entità rappresentanti le categorie di base, ma esistono alcune semplificazioni, DOLCE-Lite e DOLCE-Ultralite, che sono versioni ridotte dell'ontologia originale, più facili da integrare.

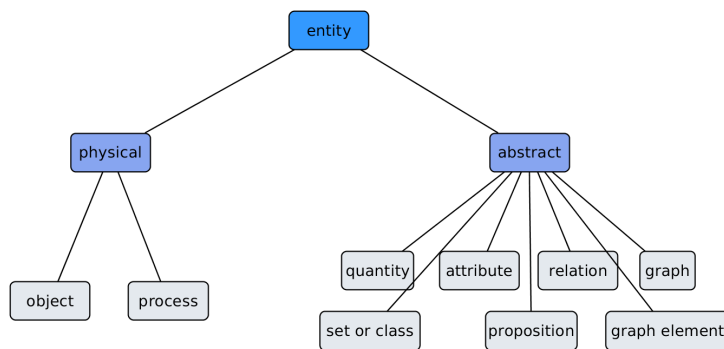
Di stampo decisamente più pragmatico è l'ontologia SUMO (Suggested Upper Merged Ontology), anch'essa *top-level*, ma realizzata combinando i contenuti di diverse ontologie disponibili in rete. In particolare sono state analizzate, classificate e mappate le entità e le relazioni di diverse ontologie in un'unica struttura. SUMO non si fonda su una particolare concezione teorica, ma è stata concepita sulla base di principi pratici, prendendo materiale già costruito e cercando di mapparlo nel modo più coerente possibile. Proprio questo fatto rende SUMO un'ontologia un po' semplicistica dal punto di vista della rappresentazione

Figura 7: La struttura di DOLCE



del mondo, ma chiara e facilmente impiegabile all'interno di algoritmi di NLP. La figura seguente mostra i primi tre livelli della gerarchia, in cui tutte le entità fanno capo a un elemento radice *entity*; il secondo livello contiene la distinzione tra entità fisiche e astratte.

Figura 8: La struttura di SUMO



2.2.4 PAROLE/SIMPLE

Attraverso il progetto europeo PAROLE è stata costituita una risorsa linguistica generale per 12 lingue europee (catalano, danese, olandese, inglese, francese, finlandese, tedesco, greco, italiano, norvegese, portoghese e svedese), per ciascuna delle quali è stato

realizzato un corpus da 20 milioni di parole e un lessico di 20.000 lemmi. I 12 lessici di PAROLE sono corredati di informazioni morfologiche e sintattiche. L'obiettivo del progetto europeo SIMPLE (Semantic Information for Multifunctional Plurilingual Lexica) è stato quello di aggiungere informazione semantica ai lessici di PAROLE. Partendo da una lista di concetti di base (100 verbi, 500 nomi, 200 aggettivi) è stato sviluppato un set di 10.000 unità semantiche, collegate alle unità sintattiche e alle alternanze rappresentate in PAROLE e codificate come un modello ontologico unificato, contenente informazioni su dominio, classe semantica, tipologia verbale, caratteri ereditari, glossa, struttura predicativa, struttura di *qualia* e polisemia (Lenci et al., 2000). La risorsa finale, nota come PAROLE/SIMPLE, è costituita da 12 lessici contenenti informazione morfologica, sintattica e semantica, organizzati in una struttura dati comune. Recentemente PAROLE/SIMPLE è diventata un'ontologia web, grazie ad una conversione in RDF e ad un mapping su *lemon* (Villegas and Bel, 2014).

2.2.5 *BabelNet*

BabelNet⁸ (Navigli and Ponzetto, 2012a) è una rete semantica multilingue, un dizionario enciclopedico strutturato in ontologia. Alla base della risorsa c'è la combinazione di due tra le più importanti basi di conoscenza, una linguistica e una enciclopedica, liberamente disponibili online: WordNet e Wikipedia. Si tratta, ad oggi, della più estesa risorsa multilingue per la disambiguazione, giunta alla versione 3.5, che ricopre 274 lingue. Informazione semantica e informazione enciclopedica sono state raccolte e collegate attraverso un algoritmo di mappatura automatica, al fine di creare un dizionario di concetti ed entità nominate caratterizzato sia da ricchezza informativa, sia da una fitta rete di rapporti semantici a livello ontologico. Concetti ed entità nominate sono rappresentati per mezzo di BabelSynset, estensione del concetto di *synset* utilizzato in WordNet. Un BabelSynset corrisponde a un concetto unitario a cui sono collegate le parole che nelle varie lingue vi si riferiscono, corredate da proprietà semantiche, una glossa ed esempi d'uso. Un ulteriore contributo giunge dalla famiglia di risorse collegate a Wikipedia, attraverso cui è stato possibile collegare ai BabelSynset le immagini

⁸ <http://babelnet.org>

archivate in Wikimedia Commons e offrire così tale informazione a supporto della disambiguazione. I numeri di BabelNet sono impressionanti: 274 lingue, oltre 80 milioni di lemmi, che si proiettano su quasi 14 milioni di BabelSynset; 40 milioni di glosse, 10 milioni di immagini, il tutto convertito in RDF per un totale di quasi 2 miliardi di triple. La versione RDF di BabelNet si basa sul modello dati di *lemon* ed è stata inserita nel LLOD Cloud (Ehrmann et al., 2014).

2.3 LE ONTOLOGIE IN CONTESTO MULTILINGUE

Un'ontologia linguistica che tratta relazioni semantiche deve scontrarsi con il problema della rappresentazione della polisemia delle parole, che è di per sé difficile da determinare e, soprattutto, da formalizzare. Il significato di una parola non è accessibile alla conoscenza umana e non è quindi determinabile senza l'analisi empirica degli usi linguistici. Per superare questo problema si può però fare affidamento a grandi corpora linguistici e su un'ampia varietà di metodologie di annotazione manuale e *data mining* che si sono sviluppate in una lunga tradizione di ricerca nell'ambito della linguistica computazionale. Attraverso l'analisi del dato empirico, infatti, un annotatore madrelingua può estrarre le occorrenze di un lemma e, grazie alla sua competenza semantica, giudicarne l'applicabilità al contesto e fornirne una classificazione. Se si ha a disposizione un corpus sufficientemente grande e vario è quindi possibile determinare l'insieme dei campi di variazione di un lemma, attraverso procedure di analisi che si basano su giudizi di parlanti madrelingua.

Più complessa è invece la formalizzazione della polisemia, poiché non sempre i sensi sono separabili e discreti, ma spesso si sovrappongono ed hanno un certo margine di vaghezza. Questo costituisce uno scoglio nella realizzazione di ontologie linguistiche, in cui i sensi di una parola sono entità separate e ben definite, che si collegano tra loro attraverso relazioni semantiche predefinite⁹. Decidere quali e quanti sono i sensi di una parola è un compito difficile, se non impossibile, tanto che in molti studi di linguistica computazionale spesso viene negata l'esistenza di un insieme assoluto di sensi e sostituito con concezioni meno rigide e più legate al contesto applicativo. Kilgarriff (1997) sostiene

⁹ Le relazioni classiche delle ontologie sono: sinonimia, iperonimia-iponimia, meronimia-olonimia, antonimia.

che i sensi di per sé non esistono, ma esistono soltanto come task relativo; Cimiano et al. (2013) riprendono quest'idea con il principio di "semantics by reference" per cui il numero di sensi di una parola dipende, in prospettiva computazionale, dal dominio di applicazione degli algoritmi di disambiguazione.

In questo contesto appare chiaro che il ruolo delle ontologie deve essere ridefinito ed in particolare quello delle ontologie interlinguistiche: in altri termini dobbiamo domandarci se e come sia possibile definire concetti discreti universali. Una strategia molto pratica è quella proposta da Resnik e Yarowsky (1999) per cui i sensi delle parole sono determinati dalle diverse traduzioni del lemma in un'altra lingua: in questo modo si ottiene una risorsa in cui i concetti non riguardano strettamente la semantica dei lemmi nelle lingue, ma vengono specificati dal differenziale in traduzione. Hovy e Nirenburg (1992) sostengono che l'unico approccio possibile per la costituzione di un inventario di sensi indipendente dalla lingua, sia quello incrementale: non si può costruire un'ontologia *language-independent*, ma ci si può avvicinare asintoticamente ad essa via via che si aggiungono nuove lingue.

Una strategia che viene spesso utilizzata per ontologie multilingui è quella di limitarsi alla rappresentazione del *core-meaning* dei lemmi, ossia di identificare soltanto i nuclei semantici più prominenti. Questo porta ad avere un ridotto numero di concetti a cui i lemmi delle varie lingue vengono legati per similarità (*near-synonyms*) e non per sinonimia, che è una relazione evidentemente troppo rigida in questo contesto. Il contributo portato da risorse che si basano su questa rappresentazione superficiale del senso è molto alto per task di NLP, poiché consente lo sviluppo di grandi reti semantiche multilingui: la suddivisione *coarse-grained* dei sensi infatti può essere realizzata in modo automatico o semi-automatico ed i sensi di parole di lingue diverse possono essere mappati insieme all'interno di un'unica entità ontologica che ne rappresenta il *core-meaning*. Una *knowledge base* come BabelNet è un esempio di risorsa di questo Tipo ed è ad oggi uno degli strumenti più utilizzati per l'elaborazione automatica del linguaggio. Naturalmente l'impatto di queste risorse per la linguistica è meno rilevante, poiché tutte quelle differenze semantiche fini che sono proprie delle parole si perdono e viene tutto ricondotto ad una rappresentazione approssimata che limita molto l'analisi del significato.

Il problema della determinazione dei sensi di parola è esplorato da vari punti di vista anche in filosofia del linguaggio. A partire da Wittgenstein l'idea di significato assume una connotazione sempre meno basata sulla determinazione di condizioni di applicabilità necessarie e sufficienti, in un'ottica più relativizzata. In particolare il significato viene considerato strettamente legato all'uso linguistico: *the meaning of a word is its use in the language* (Wittgenstein, 1953). Il Generative Lexicon di Pustejosky (1991), costruito su base empirica attraverso l'analisi dei corpora, si fonda sull'idea che i sensi non esistano senza i contesti. Una posizione ancora più radicale la si trova in Croft e Cruse (2004): il senso non esiste di per sé, ma è creato nel momento stesso in cui viene usata la parola. In queste diverse concezioni teoriche si rileva una linea comune che mette in stretta correlazione il senso con gli usi linguistici e che è molto lontana da un'idea fregeiana di senso come entità oggettiva.

Da queste osservazioni emerge una certa incompatibilità tra la rappresentazione dei sensi di parola in un'ontologia formale per il trattamento automatico del linguaggio, che prevede una rappresentazione semplificata dei concetti in entità discrete e separabili, e una risorsa linguistica basata su un'analisi fine del senso. IMAGACT rappresenta un caso abbastanza particolare nel panorama delle ontologie linguistiche: essa infatti si basa su una procedura di identificazione dei concetti su base pragmatica (in accordo con una concezione wittgensteiniana), grazie alla quale riesce a fornire una categorizzazione fine dell'estensione di un lemma e, parallelamente, una rappresentazione discreta delle entità concettuali in contesto multilingue. Pertanto IMAGACT può essere sfruttata sia per condurre analisi linguistiche sul senso e sul riferimento, sia come supporto di conoscenza in algoritmi di disambiguazione e traduzione.

L'ONTOLOGIA DELL'AZIONE IMAGACT

3.1 CONCETTI DI BASE

3.1.1 *I linguaggi naturali e l'ontologia dell'azione: i verbi generali, l'equivalenza locale e la produttività nelle traduzioni*

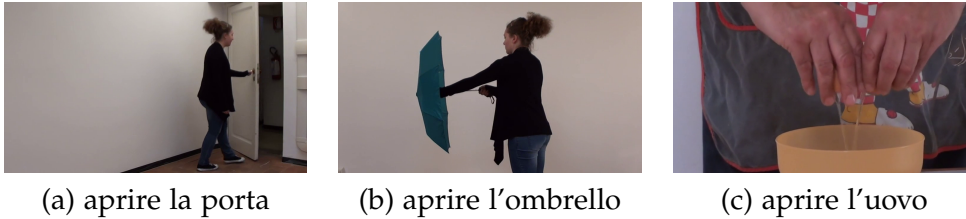
Un'azione è un modello di modificazione del mondo da parte di un attore che può essere applicato a un insieme aperto di oggetti e ogni oggetto può sottostare a un insieme aperto di azioni. Ciò significa che le azioni sono entità ontologiche produttive. Lavori classici di psicologia, filosofia e intelligenza artificiale concordano su questa proprietà delle azioni (von Wright, 1963; Minsky, 1969; Tomasello, 2003). La produttività dell'azione è rispecchiata a livello linguistico. Un predicato di azione può essere applicato ad un insieme aperto di argomenti che possono essere anche argomenti di azioni diverse. La capacità di fare atti di predicazione emerge nell'ontogenesi quando entrambe le condizioni sono soddisfatte (Piaget, 1954; Halliday, 1975; Bruner, 1983).

Ma i predicati delle lingue naturali vanno aldilà di questo livello di produttività. Per esempio, proprio come nel caso di *girare* e *turn*, presentati nell'introduzione, l'istruzione inglese *open* può portare a azioni qualitativamente diverse:

- (a) opening a door
- (b) opening the umbrella

In altre lingue, come le lingue romanze, e l'italiano in particolare, la variazione è anche maggiore e il verbo corrispondente (*aprire*) si può estendere anche ad eventi simili all'evento rappresentato in figura 1c (*aprire l'uovo*).

Nelle circostanze precedenti ha luogo più di un Tipo di azione e non è ovvio quali proprietà gli eventi in (a), (b) e (c) condividono per essere individuabili con un identico

Figura 9: Tre tipi di azione diversi per il verbo *aprire*

predicato. Per esempio l'azione di aprire la porta permette l'accesso allo spazio esterno, mentre nel caso dell'ombrello non c'è spazio esterno ma il riferimento a proprietà funzionali dell'oggetto; in aprire l'uovo l'accesso allo spazio interno e al suo contenuto è realizzato con deterioramento dell'oggetto. In altre parole ciò che il linguaggio designa come una singola azione è riferito a una variazione di tipi azionali diversi. Questo giudizio è confermato dalla produttività di ogni Tipo. Per esempio siamo capaci di giudicare che al contrario lo stesso Tipo di azione si realizza in ognuna delle serie in (a), (b) e (c) nonostante che gli oggetti implicati siano diversi:

- (a) Bill ha aperto la porta / la finestra / il recinto / la tenda
- (b) Bill ha aperto l'ombrello / la sedia a sdraio / la biro / il lucchetto
- (c) Mario ha aperto la noce / l'uovo / il pacco

Chiamiamo "generali" i verbi che condividono questa proprietà.

Wittgenstein (1953) con la speculazione sulla variazione del concetto di *gioco* costituisce probabilmente il riferimento teorico più antico per l'identificazione della variazione pragmatica di un predicato, ma la nozione è stata messa in evidenza negli anni '90 da Moneglia per la ricerca semantica e da Bowerman e associati per gli studi acquisizionali (Moneglia, 1987; Choi and Bowerman, 1991); recentemente la nozione è stata riconsiderata estensivamente negli studi tipologici (Kopecka and Narasimhan, 2012).

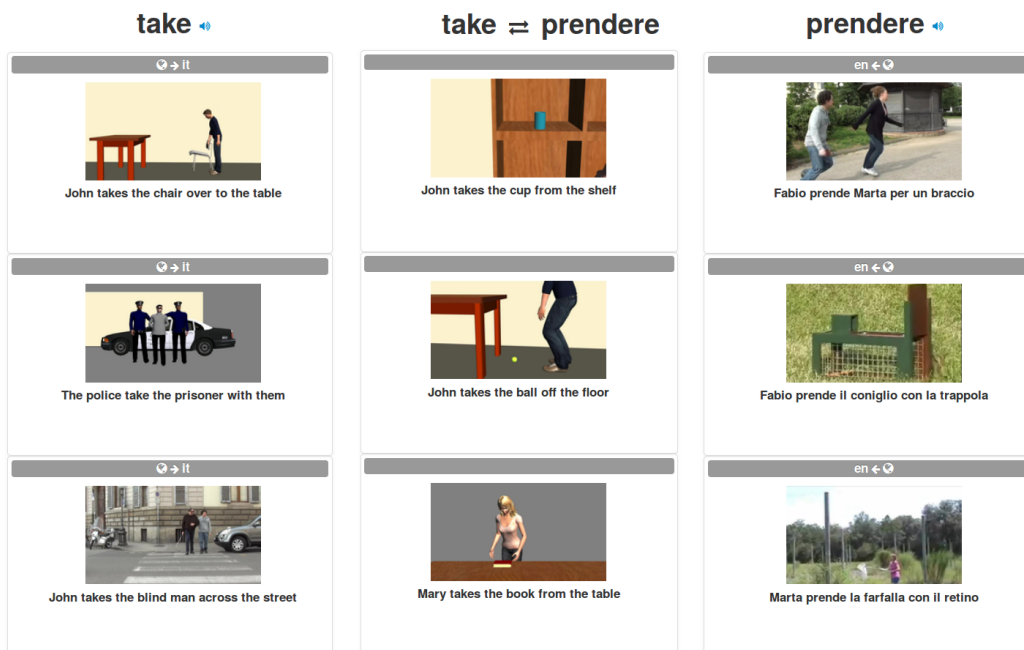
Nel caso dei verbi generali il linguaggio ordinario non rispecchia dunque l'ontologia dell'azione e causa un notevole problema di Natural Language Understanding. Di fatto, il *label* linguistico non specifica l'entità ontologica a cui si riferisce. Il problema diviene anche più sensibile a livello cross-linguistico, dato che i verbi generali rendono problematico il riferimento interlinguistico alle azioni anche più comuni. Possiamo replicare il caso già osservato per girare e aprire anche per il verbo di azione a più

3.1 CONCETTI DI BASE

alta frequenza in italiano, *prendere*, che corrisponde a *take* in inglese. La variazione ontologica essenziale dei due verbi, esemplificata dai seguenti concetti prototipali in IMAGACT, lo mostra chiaramente (Fig. 10):

- i due verbi sono generali perché si estendono a eventi qualitativamente diversi;
- ciascun Tipo si caratterizza per il fatto di poter essere identificato equivalentemente da altri verbi (localmente equivalenti);
- nonostante *prendere* e *take* siano, in teoria, in relazione di traduzione, la loro applicazione corre in parallelo solo in un sottoinsieme dei tipi (al centro), ovvero nei tipi rispettivamente equivalenti a *togliere*, *rubare* e *dare*;
- *take*, ma non *prendere*, si applica a tipi di azione in cui sono equivalenti *bring*, *lead* (sulla sinistra); *prendere*, ma non *take*, si applica a tipi equivalenti a *acchiappare* (sulla destra).

Figura 10: Comparazione *prendere* / *take* e tipi differenziali



Non esiste dunque corrispondenza uno a uno tra *prendere* e *take*, dato che non sono equivalenti rispetto alle loro possibili estensioni e quindi non sono equivalenti in intensione. Come si diceva, in particolare, *prendere* non si estende ai campi in cui *take* è equivalente a

lead o *bring* (sulla sinistra). In quei casi l'italiano richiede *portare* o *accompagnare*, mentre *take* non si estende ai campi in cui *prendere* è equivalente ad *acchiappare* (sulla destra). In quel caso l'inglese richiede *catch*. Tali tipi di azione sono dunque differenziali.

Il precedente insieme di corrispondenze è una sorgente estremamente ricca di informazioni per quanto riguarda la relazione tra lemmi e tipi azionali:

- ogni verbo si applica nel suo significato proprio a vari tipi azionali e nessuno di questi può essere considerato più appropriato dell'altro per caratterizzarne il significato (nessuno è più prototipico dell'altro);
- i verbi sono produttivi in ogni Tipo e anche la relazione di traduzione appare produttiva all'interno del Tipo.

Per esempio, se si considerano inglese e italiano, dato che *take* e *prendere* sono in relazione di traduzione nei tipi al centro e *take* e *portare* sono in relazione di traduzione nei tipi a sinistra, si può prevedere che la relazione di traduzione sussista in tutte le istanze di ogni Tipo, indipendentemente dall'oggetto che è coinvolto nell'azione, come nelle serie seguenti:

- l'attore <*takes / prende*> <un libro / una penna / una tazza / una bambola> <dallo scaffale / dal mobiletto / dalla tavola>
- l'attore <*takes / porta*> <il cieco/ il cavallo/ il ladro> da qualche parte.

Se si considera la produttività dei tipi azionali che caratterizzano la variazione di significato dei verbi generali si può dedurre che la presenza della caratterizzazione delle azioni attraverso questi verbi non dipende dalla fraseologia idiosincratICA delle lingue, ma è un aspetto della competenza semantica di ogni lingua.

3.1.2 *La produttività dei concetti: estensioni primarie e marcate dei verbi: Identificazione e definizione*

La presenza del problema di disambiguazione e traduzione dei verbi di azione generali non è solo teorico ma riguarda l'uso linguistico quotidiano: la maggior parte delle volte in cui ci si riferisce ad una azione lo si fa con un verbo generale, ovvero le azioni della vita

ordinaria sono molto spesso categorizzate linguisticamente con un verbo generale, anche se lingue diverse possono riscontrare una maggiore o minore tendenza ad utilizzare verbi generali o più specifici. Dato però che l'applicazione di tali verbi ai tipi azionali nella loro estensione è produttiva, allora questa dovrebbe essere, in linea di principio, anche predicibile e quindi sfruttabile in sistemi di disambiguazione automatica delle lingue naturali. La necessità di disambiguazione del riferimento all'azione è ovvia sia per task di traduzione che per la comunicazione uomo macchina: il linguaggio non specifica l'azione a cui si riferisce tra le varie estendibili da un verbo generale e non c'è garanzia di traduzione. Il risultato principale di IMAGACT è l'identificazione delle entità ontologiche produttive nel dominio dell'azione che sono specificate dai verbi di azione più frequenti in italiano e in inglese e l'estensione delle relazioni tra verbi e entità ontologiche al cinese e allo spagnolo. L'ontologia fornisce la sorgente di informazione per l'interpretazione del linguaggio naturale e quindi per la sua disambiguazione e traduzione automatica, permettendo di fondare il design di sistemi che interagiscono in contesto naturale attraverso istruzioni in linguaggio naturale. La distinzione tra variazioni produttive e non produttive di un verbo è quindi essenziale per determinare previsioni basate su criteri semantici.

Le risorse esistenti, e in particolare WordNet, che costituisce la principale base di dati lessicale oggi disponibile, non contengono informazione sufficiente a questo scopo per una varietà di ragioni (Moneglia et al., 2012a). In particolare, i concetti elencati per la disambiguazione di un verbo rendono poco prevedibile la possibilità di traduzione, ovvero la produttività dell'applicazione del verbo non può essere garantita da tutti i *synset* di WordNet nella stessa misura. I verbi hanno infatti vari usi che si distaccano dal loro significato effettivo, ed in questi significati la relazione di traduzione non può essere predetta. Ad esempio, tra i *synset* di WordNet del verbo *to put* è riportato il seguente:

S: (v) arrange, set up, put, order (arrange thoughts, ideas, temporal events)

In questa entrata dell'ontologia, diversamente da quanto avviene nei casi precedenti, la possibilità di traduzione non corre in parallelo in tutte le istanze potenziali del Tipo. Funziona in (1), ma per qualche ragione idiosincratICA, non in (2):

(1) I put my schedule in a certain way > Ho messo i miei impegni in un certo modo

(2) I put my life in a certain way > * Ho messo la mia vita in un certo modo.

La distinzione tra tipi produttivi e tipi idiosincratici è cruciale: solo gli usi primari sono sicuramente produttivi, mentre gli usi fraseologici o metaforici, nella maggior parte dei casi, non lo sono. In altri termini, mentre la variazione in figura 10 identifica le variazioni in estensione su tipi di azioni diverse che un parlante nativo deve poter assentire o rifiutare sulla base della sua sola competenza linguistica, lo stesso non vale per usi marcati come in (1). Solo l'identificazione degli usi produttivi costituisce una base di conoscenza per la previsione degli ambiti di estensione dei verbi di lingue diverse nello spazio dell'azione e per rendere obiettive le relazioni di traduzione. Da questo punto di vista una ontologia dovrebbe distinguere chiaramente tra i primi, che costituiscono prototipi semantici e i secondi, che costituiscono usi idiosincratici e non sono quindi concetti dello stesso Tipo.

Crucialmente solo l'identificazione, e non la definizione delle entità individuate, è richiesta per stabilire le relazioni cross-linguistiche. Nei termini wittgensteiniani: *come posso spiegare a qualcuno cos'è un gioco? Semplicemente indicando un gioco e dicendo "Questo e simili cose sono giochi"* (Wittgenstein, 1953). Lo scenario "alla Wittgenstein" è utilizzato in IMAGACT sia per distinguere le variazioni produttive dalle variazioni non produttive all'interno dell'uso linguistico dei verbi, sia per identificare tipi azionali a livello cross-linguistico, consentendo la comparazione diretta dei tipi derivati dall'annotazione dei corpora di lingue diverse. Il primo compito sfrutta lo scenario "alla Wittgenstein" come banco di prova della effettiva produttività dei concetti. Si deve notare, infatti, che solo gli usi che ad un parlante competente appaiono adeguati a rappresentare il significato di un predicato possono essere indicati come prototipi per l'uso del predicato stesso. In parallelo, gli usi non primari o comunque metaforici o fraseologici non possono essere indicati come istanze prototipiche di ciò che viene significato.

Si consideri ad esempio il verbo italiano *rotolare*. L'istanza (3), derivata da corpus, può essere ragionevolmente indicata come una istanza prototipica del concetto espresso dal verbo. In altri termini un parlante competente può indicare questa istanza a qualcuno che non conosce la lingua fornendo l'informazione sufficiente a specificarne il significato con

la seguente locuzione: “questa e simili cose sono ciò che noi intendiamo con rotolare”. Al contrario, l’istanza (4) non potrà ragionevolmente essere indicata come un’istanza di “ciò che noi intendiamo con rotolare”.

(3) Cristina si rotola nell’erba umida

(4) Il bambino rotolò in terra dal seggiolone

Infatti, nonostante la frequenza con cui può comparire in quel contesto, in (4) il verbo è usato palesemente in senso non proprio (il bambino non rotola, bensì cade) e ciò risulta evidente ad un parlante competente.

L’indicabilità di una istanza come istanza campione può essere usata come un test della competenza che consente, salvo casi limite, di isolare la gran parte degli usi strettamente propri del verbo, identificando poi la loro variazione¹. Lo stesso avverrà con le frasi derivate dal corpus inglese. Ad esempio, per quanto riguarda la variazione del verbo *roll*, (5) potrà essere indicata come un istanza prototipica di ciò che si intende con *to roll*, ma non (6).

(5) John rolls a cigarette

(6) John rolls the words around in his mind

Lo studio della variazione produttiva di un verbo inizia quindi quando la variazione propria di un predicato è distinta dalla variazione di significato metaforica e fraseologica. Il progetto IMAGACT ha avuto come obiettivo di identificare le variazioni significative dei verbi di azione su una ontologia dell’azione di più lingue ad alto impatto nella società contemporanea (in particolare italiano, inglese, cinese e spagnolo) garantendo le relazioni di traducibilità tra i predicati delle lingue che si applicano a ciascun concetto azionale.

¹ Vedi oltre per le procedure di estrazione dei concetti da corpus in IMAGACT e il capitolo 8 per la validazione di questo principio

3.2 IL FRAMEWORK DI IMAGACT

3.2.1 *La Strategia corpus-based*

IMAGACT tratta la proprietà centrale del lessico verbale delle lingue naturali relativamente al modo con cui queste esprimono l'azione. La maggior parte dei verbi di azione ad alta frequenza, infatti sono "Generalisti", ovvero si estendono, nel loro significato proprio, a classi di azioni fisiche tra loro assai diverse e risultano quindi intrinsecamente ambigui (non specificano l'azione a cui si riferiscono, bensì più azioni). La tabella seguente indica i verbi di azione più alti in frequenza nei corpora italiano e inglese e il numero di tipi azionali riscontrati per ognuno di essi.

Tabella 2: Verbi d'azione più alti in frequenza nei corpora di parlato italiano e inglese

Verbo	Frequenza	Tipi	Verbo	Frequenza	Tipi
mettere	4018	16	take	4006	13
dare	3261	3	put	3231	13
prendere	2459	16	give	3030	3
portare	1509	8	keep	1282	7
passare	946	6	leave	1272	5
lasciare	914	9	bring	851	7

Inoltre, lingue diverse si riferiscono all'azione con loro predicati generali che risultano solo parzialmente equivalenti, rendendo difficile sia la traduzione automatica della frase semplice sia l'acquisizione in L2 di una parte centrale della competenza semantica delle lingue oggetto di acquisizione.

Nonostante la centralità di questo fenomeno e il suo impatto per l'elaborazione, sia pratica che formale, dell'informazione linguistica naturale, le equivalenze e i differenziali degli ambiti di riferimento dei verbi generali, non sono derivabili con sicurezza dalle infrastrutture linguistiche disponibili (dizionari, dizionari bilingui, ontologie quali WordNet, FrameNet, VerbNet, PropBank ecc.). Questo è vero anche rispetto alle lingue con maggior tradizione lessicografica e a maggior ragione per l'insieme delle lingue che vengono a contatto nelle società multiculturali. In altri termini la conoscenza dell'ambito di variazione effettivo dei verbi d'azione in lingue diverse costituisce un problema empirico complesso. Questo emerge chiaramente dai test effettuati su BabelNet (vedi oltre al

capitolo 8), nel quale si evidenziano da un lato una rappresentazione incompleta della variazione dei verbi e dall'altro un certo grado di vaghezza e sovrapposizione tra alcuni sensi. Ultimamente i processi di disambiguazione sono compromessi, come vedremo nel capitolo 8, in una percentuale molto alta (circa l'80% dei casi dall'inglese all'italiano e circa il 50% dall'italiano all'inglese, sul dataset di *testing*).

D'altro canto la conoscenza su quale sia effettivamente l'universo delle azioni rilevanti nell'orizzonte naturale è tutt'altro che a disposizione, mentre al contrario la modellazione dell'azione è attualmente un tema centrale, oltre che in ambito linguistico e cognitivo, anche in robotica e in neurologia (Borghini and Cangelosi (2014); Cangelosi and Schlesinger (2015); Veskos and Pastra (2014); van Dam et al. (2010)). Non esistono in particolare basi di dati relative all'universo azionale naturale.

IMAGACT colma questa lacuna, almeno rispetto alle azioni riferite con maggior frequenza nel lessico dei verbi d'azione, e relativamente a un numero di lingue, limitato ma ad alto impatto a livello sia nazionale che globale: italiano e inglese e, a partire da queste, il cinese mandarino e lo spagnolo. La metodologia utilizzata per identificare le azioni sfrutta i corpora di parlato spontaneo e basa su quelli un processo induttivo. Il riferimento linguistico alle azioni e la rilevanza di tali azioni nella nostra vita vanno infatti di pari passo. L'uso reale dei verbi di azione nella performance linguistica può essere apprezzato osservando le occorrenze di tali verbi nei corpora in cui i parlanti si riferiscono alle azioni. I corpora documentano come le azioni sono rappresentate nel linguaggio e possono quindi essere sfruttati per estrarre questa informazione. Dato che nel parlato, rispetto alla comunicazione scritta, si concentra sia un maggior numero di verbi, sia la probabilità del riferimento ad azioni fisiche rilevanti nella vita quotidiana, e dato che sono attualmente a disposizione risorse significative per la rappresentazione del parlato, l'informazione sull'ontologia dell'azione e sul modo in cui lingue diverse si riferiscono ad essa è stata estratta attraverso un processo di annotazione di corpora di parlato spontaneo (Moneglia and Panunzi, 2007).

IMAGACT ha derivato da corpora di parlato spontaneo italiani ed inglesi (il BNC e una collezione di corpora italiani) l'informazione sulla varietà dei diversi tipi di azioni riferiti nel loro significato proprio dai verbi generali ad alta frequenza e più estesamente dai verbi di azione che si riferiscono all'azione fisica (Moneglia et al., 2012b, 2014a). Si

deve sottolineare che la variazione di significato dei verbi considerata in IMAGACT solo rispetto ai concetti in cui il verbo si estende in senso proprio ed esclude i significati traslati e metaforici, limitandosi quindi alle sole azioni fisiche. IMAGACT specifica infatti solo le azioni che costituiscono libere estensioni dal verbo, accessibili ai parlanti competenti senza specifiche restrizioni d'uso, ovvero specifica i concetti rispetto ai quali l'uso è produttivo e come abbiamo visto, tale previsione, nel dominio del riferimento all'azione, non può riguardare i significati traslati dei verbi, che sono stati identificati nei corpora e archiviati. La procedura per la determinazione del Tipo di variazione, primaria o marcata, di un verbo d'azione non richiede particolari conoscenze, se non la competenza semantica della lingua. Un primo criterio discriminativo è dato dal fatto che il verbo all'interno della frase si riferisca ad un'azione fisica: questo infatti costituisce un requisito per poter avere variazione primaria, dal momento che l'ambito d'analisi è circoscritto ai verbi d'azione.

(7) Marco prende il bicchiere (var. primaria)

(8) Marco prende l'iniziativa (var. marcata)

In (8) ad esempio il verbo prendere è applicato ad un processo mentale e fa quindi parte di un ambito diverso da quello strettamente fisico.

Sebbene il riferirsi ad un'azione fisica sia un tratto necessario per avere un'istanza primaria del verbo, questo non è sufficiente: ci sono infatti usi marcati che predicano di azioni fisiche. Gli esempi (9) e (10) descrivono eventi in cui c'è un'azione fisica, peraltro descritta attraverso applicazioni molto comuni del verbo *prendere*, ma fanno parte della variazione marcata, in quanto il verbo non è utilizzato nel suo senso "proprio".

(9) Marco prende un caffè

(10) Marco prende l'autobus

Per quanto detto nel paragrafo precedente, per determinare con certezza se un'istanza fa parte della variazione primaria o meno, si utilizza il test di Wittgenstein che, in una versione adattata al nostro ambito di applicazione, può essere espresso come segue:

è possibile dire a qualcuno che non conosce il significato del verbo x che "l'azione a cui il verbo si riferisce ed azioni simili sono ciò che intendiamo con x"?

In caso di risposta affermativa abbiamo un uso primario del verbo, altrimenti un uso marcato.

Provando ad applicare il test ai casi (9) e (10) notiamo che, da parlanti italiani competenti, rispondiamo negativamente: se davanti all'azione di "prendere un bicchiere" (7) possiamo asserire con certezza che quella e simili cose sono ciò che intendiamo con prendere, non possiamo dire altrettanto per azioni come "prendere un caffè" o "prendere un autobus". Queste e simili cose infatti non sono ciò che intendiamo con *prendere*, ma ciò che intendiamo con *bere* (9) o con *montare* (10). In generale il task ha un alto livello di *inter-rater agreement*: l'esperimento di Gagliardi (2014) sulla distinzione tra occorrenze verbali primarie e marcate misura una *k* di Cohen di 0.96 (vedi più oltre al capitolo 8). Soltanto gli usi primari sono stati oggetto di analisi e annotazione linguistica, che ha portato ad una categorizzazione degli ambiti di applicazione dei verbi d'azione ed alla definizione dei concetti pragmatici che ne rappresentano l'estensione.

IMAGACT specifica la variazione di significato dei verbi utilizzando il linguaggio universale delle immagini e non attraverso definizioni. Ogni concetto azionale costituisce un oggetto ontologico indipendente, identificato attraverso una scena prototipica alla quale sono poi associati i verbi che possono estendere tale concetto in senso proprio. Dato che la scena è un oggetto identificabile da parte di un utente umano, come si diceva in precedenza, l'ordine di presentazione dell'informazione in IMAGACT può essere diverso dal dizionario, che ordina i termini alfabeticamente e ne dà definizione².

3.2.2 *Il processo induttivo da corpus*

L'informazione ontologica in IMAGACT, è basata sulla annotazione di più di 100.000 occorrenze di verbi d'azione ad alta probabilità di occorrenza nei due corpora italiani e inglesi scelti come dataset di partenza. Tale dataset multilingue costituisce la base per la creazione di una ontologia dell'azione indipendente dal linguaggio, ovvero una ontologia che tende a definire la varietà di azioni riscontrate nell'uso dei lemmi verbali italiani e inglesi definite indipendentemente da tali lemmi (*mapping*). La metodologia

² Al capitolo 8 sarà riportato per esteso l'esperimento in Gagliardi 2014 che garantisce l'identificabilità dei tipi d'azione espressi dalle frasi attraverso scene.

di induzione è messa in opera attraverso una infrastruttura di annotazione di rete e ha seguito il seguente schema³:

1. standardizzazione delle occorrenze nel corpus e selezione dei contesti in cui il verbo è usato in senso proprio;
2. generazione dei tipi di azioni e raggruppamento delle istanze sotto *best example* per ogni Tipo;
3. annotazione dei *best example* dei tipi con struttura tematica e aspettuale, validazione delle occorrenze nel Tipo e simultaneo allineamento di tutte le standardizzazioni rispetto alla struttura tematica.

A partire dai riferimenti ad azioni di un verbo d'azione nel suo senso proprio, un parlante competente è in grado di raggruppare queste azioni per similarità. Tra i riferimenti di un verbo d'azione in un corpus troviamo infatti eventi pragmaticamente simili ed altri invece molto diversi. Il verbo *girare*, che qui prendiamo come esempio di verbo generale, ha un'ampia variazione e può predicare di eventi difficilmente assimilabili per tratti pragmatici, come "Il tassista gira in via Cavour" e "Il cuoco gira la zuppa". Nonostante che queste frasi contengano entrambe un uso primario del verbo *girare*, non c'è vicinanza pragmatica tra gli eventi a cui esse fanno riferimento. Viceversa la frase "Marco gira a destra" può essere facilmente giudicata simile a "Il tassista gira in via Cavour"; benché gli eventi denotati siano diversi, infatti, essi hanno una serie di caratteri comuni che consentono ad un parlante di raggrupparle all'interno di un unico concetto azionale. Le frasi seguenti esemplificano il compito dell'annotatore nel processo di tipizzazione delle occorrenze:

(11) Il croupier gira la roulette

(12) La trottola gira

(13) Marco gira le zucchine

(14) Il cuoco gira la zuppa

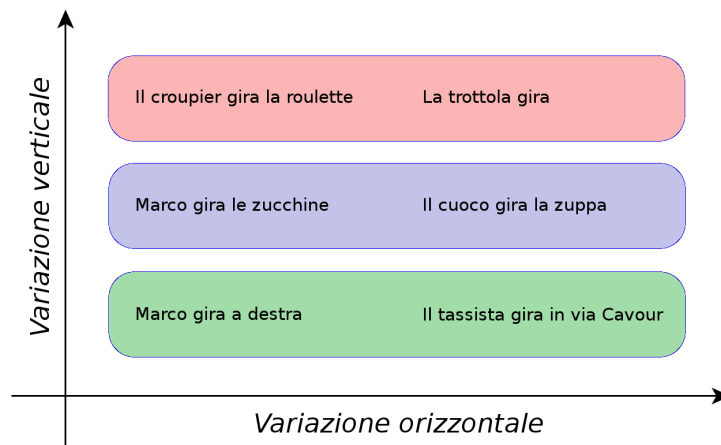
³ Il processo di annotazione che porta alla identificazione dei tipi di azione da corpus è stato pubblicato in più occasioni; si veda in particolare la presentazione all'ISO standard per l'annotazione semantica (Moneglia et al. (2012a))

(15) Marco gira a destra

(16) Il tassista gira in via Cavour

In queste sei frasi possiamo vedere abbastanza chiaramente la similarità tra le coppie 11-12, 13-14 e 15-16, che evidenziano ciascuna un concetto pragmatico all'interno della variazione di *girare*; altrettanto evidente è il fatto che qualsiasi altra ipotesi di raggruppamento di queste 6 frasi sia inappropriata. All'interno della variazione primaria di un verbo distinguiamo quindi due sotto-tipi di variazione: una variazione verticale, che determina l'appartenenza delle azioni a diversi concetti azionali, ed una orizzontale, che rappresenta la varietà di azioni categorizzabili all'interno di uno stesso concetto.

Figura 11: Variazione verticale e orizzontale



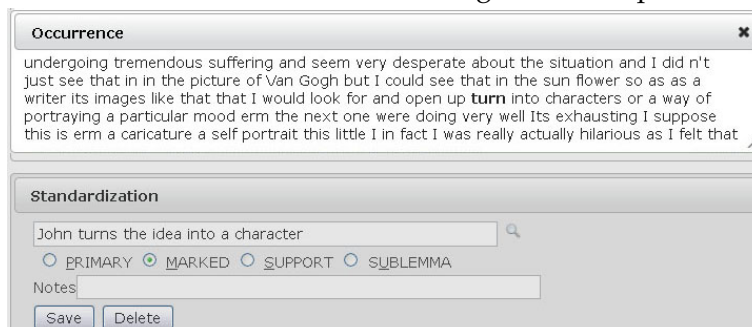
La variazione verticale definisce i Tipi azionali derivati da una annotazione. E' importante qui evidenziare la differenza teorica tra un Tipo e un concetto in una ontologia dell'azione: quest'ultimo infatti dovrebbe essere indipendente dalla lingua, mentre il Tipo è l'induzione di una classe di azioni simili nell'ambito di variazione verticale di un verbo. Il Tipo è quindi una proprietà del verbo, mentre il concetto azionale, in linea di principio è una entità pragmatica indipendente dalla lingua⁴.

La serie di immagini seguenti mostra la sequenza di azioni compiute per l'annotazione del verbo inglese *turn* e la selezione delle sue variazioni verticali. Ogni occorrenza viene visualizzata e standardizzata in una frase semplice attiva in cui il verbo è saturato e alla

⁴ Indipendente dalla lingua non significa universalmente condiviso: nel capitolo 5 infatti sono mostrati alcuni esempi in cui dei concetti azionali (derivati da corpora italiani e inglesi) non sono riconosciuti in cinese.

frase viene assegnato in particolare il valore di uso proprio oppure marcato. La figura 12 mostra una occorrenza marcata e quindi non oggetto di ulteriore annotazione.

Figura 12: Standardizzazione e selezione degli elementi primari e marcati



La Figura 13 mostra le occorrenze giudicate primarie (in verde) che sono raggruppate nei tipi indotti dalla variazione di *turn*. I Tipi (in azzurro) sono caratterizzati da uno o più *best example* e da uno o più verbi equivalenti (in chiaro in basso). L'annotatore forma i tipi e identifica i *best example* in cui il Tipo si manifesta con diversa struttura argomentale. Al termine di questo lavoro l'annotatore classifica tutte le occorrenze standardizzate.

La figura 14 mostra l'assegnazione di struttura tematica ai *best example* (al centro). In questo lavoro di analisi fine l'annotatore allinea a questa struttura tutte le occorrenze del Tipo che vi si conformano e ne verifica la congruità. Per ogni *best example* egli assegna inoltre una classe aspettuale. L'annotatore può anche creare occorrenze *competence based* nel caso il corpus non contenga occorrenze per soddisfare la variazione che si dimostra possibile nella lingua.

Tale lavoro è stato svolto in parallelo da linguisti madrelingua italiani e inglesi.

3.2.3 *Il mapping dei tipi su una ontologia interlinguistica dell'azione e la rappresentazione per prototipi*

I Tipi Azionali possono essere considerati un livello di rappresentazione dell'azione che individua, secondo l'intenzionalità umana lingua specifica, quale sia l'oggetto rilevante di riferimento per gli umani nel dominio dell'azione. La metodologia utilizzata in IMAGACT per l'oggettivazione di tali entità è fortemente innovativa e separa l'identificazione positiva dei tipi azionali dalla loro definizione. Le definizioni sono coppie sinonimiche

Figura 13: Formazione dei tipi e *clustering* delle occorrenze nei tipi

The screenshot displays the IMAGACT interface for type formation and clustering. On the left, a vertical list shows nine types (Type 1 to Type 9) with their corresponding examples. The control panel at the top right includes buttons for 'Show all', 'Show primary not assigned to type', 'Show deleted', 'Show not primary', 'Show sublemma', and 'Jump to last modified'. Below these buttons is a pagination control showing 'Page: 1 / 5' with 'First', 'Prev.', 'Next', and 'Last' links. The middle right section contains a table of occurrences with columns for 'Left Context', 'Verb', 'Right Context', 'Status', and 'Deleted'. The bottom right section, titled 'All Equivalent Verbs used for Verb: turn', contains a table mapping various verbs to their respective types.

Left Context	Verb	Right Context	Status	Deleted
John turns right			Type 3.1	
John turns left			Type 3.1	
John turns right			Type 3.1	
John turns eastbound			Type 3.1	
The boat turns around			Type 3.1	
John turns around			Type 3.1	

All Equivalent Verbs used for Verb: turn			
to direct	Type 3	to spin	Type 2
to flip	Type 1	to spin	Type 7
to fold	Type 6	to spin	Type 9
to orientate oneself	Type 8	to stir	Type 5
to rotate	Type 4	to twist	Type 8

che dovrebbero specificare attraverso una serie di condizioni necessarie e sufficienti quali entità ricadono sotto quel concetto, permettendo nel contempo di spiegare le relazioni gerarchiche tra entrate. Come la tradizione di semantica formale del secolo scorso ha mostrato, queste definizioni sono però sia incomplete che radicalmente sottodeterminate (Quine, 1959; Peruzzi, 1982) e questo causa una sostanziale inadeguatezza a selezionare nel mondo gli oggetti a cui si riferiscono i predicati verbali di azione delle lingue naturali.

IMAGACT distingue "definizione" e "identificazione" come due compiti separati. Sfruttando la capacità dei soggetti umani di apprezzare somiglianze tra eventi in uno scenario alla Wittgenstein. Una volta operata la generazione dei tipi di azione riferiti dai verbi di azione nei due corpora italiano e inglese la formazione dell'ontologia interlinguistica prevede la mappatura di tali tipi su una serie unica di concetti azionali che costituisce l'ontologia di riferimento per entrambe le lingue. Tale lavoro consiste nella identificazione dei tipi uguali denotati dai verbi rispettivamente italiani e inglesi e parallelamente nella "riconciliazione" delle annotazioni in un'unica serie di concetti interlinguistici. La

Figura 14: Validazione, annotazione del Tipo e allineamento delle occorrenze alla struttura tematica

The screenshot displays the IMAGACT interface. On the left, a sidebar lists 'Action Types' with categories 1 through 5. Type 5 is highlighted in red, corresponding to the example 'BE1 John turns the mixture to stir'. The main panel shows the configuration for 'Type 5', including a 'Script' section with the example sentence and a 'Thematic grid' table. Below this is a 'Standardized Occurrences' table with columns for Type-BE, Standardization, Validity, Move to, Peripheral, and Actions.

Type - BE	Standardization	Valid.	Move to	Peripheral	Actions
T: 5 - BE: 1	{John}A⊗ {turns}VE {the mixture}TH	<input checked="" type="checkbox"/>	PRIMARY	<input type="checkbox"/>	[Icons]
T: 5 - BE: 1	{John}A⊗ {turns}VE {the soup}TH	<input checked="" type="checkbox"/>	PRIMARY	<input type="checkbox"/>	[Icons]
T: 5 - BE: 1	{John}A⊗ {turns}VE {the stew}TH	<input checked="" type="checkbox"/>	PRIMARY	<input type="checkbox"/>	[Icons]

metodologia sviluppata in IMAGACT per l'induzione da corpora dei tipi azionali e per la loro identificazione nell'ontologia è fondata:

- sull'uso dell'immagine (scene), piuttosto che su definizioni esplicite;
- sullo sfruttamento della proprietà della equivalenza locale.

Sono mappati sullo stesso oggetto ontologico azionale costituito da un prototipo realizzato in una scena i tipi indotti autonomamente dai corpora italiano e inglese che:

- corrispondono alla estensione di uno stesso prototipo (condizione cognitiva);
- sono identificabili dalla stessa serie di verbi equivalenti (condizione linguistica).

Dati tali criteri il mapping dei tipi derivati da corpus su una sola ontologia interlinguistica segue il requisito elementare che due tipi equivalenti non siano identificati da due "scene" diverse. L'utilizzo delle scene prototipiche, anziché delle usuali definizioni, segue da motivazioni pratiche, oltre che teoriche. Ad esempio, è facile che due soggetti parlanti lingue diverse si accordino sul fatto che il Tipo di evento corrispondente a "girare a destra" è di ordine diverso rispetto a quello rappresentato da "girare la carta" e dunque rilevare che questi tipi debbono rappresentare due concetti diversi in una ontologia dell'azione.

Questa forma di discriminazione è possibile da parte di due parlanti competenti italiani e inglesi, che possono facilmente concordare da un lato che il *best example* di un Tipo di *turn* “turn the postcard” e il *best example* di un Tipo di *girare* “girare il piatto” corrispondono allo stesso Tipo di evento diverso dal Tipo di evento istanziato dai tipi “the car turns to the left” e “il cavallo gira a sinistra”. Crucialmente, questo accordo si basa sull’individuazione e non sulla definizione delle caratteristiche di un Tipo di evento e può avvenire che anche se i tratti essenziali di ciascun Tipo rimangono indefiniti. Al contrario la definizione di quali siano le proprietà caratterizzanti i due tipi di eventi in questione richiede una teoria ed è sottoposta ad una procedura arbitraria di selezione di proprietà essenziali, che da un lato risulta sottodeterminata e dall’altro è difficile da comparare interlinguisticamente e anche intersoggettivamente: poiché infatti ogni entrata è definita in modo indipendente da due annotatori diversi, accordarsi sull’equivalenza di due definizioni è compito complesso e ultimamente improprio.

L’ontologia IMAGACT, sulla base delle evidenze fornite dalle istanze di azione reperite in un corpus, identifica quindi i tipi di azione nei termini di scene prototipiche piuttosto che attraverso definizioni, esattamente come avviene nelle figure di *girare* già viste. Le scene sono indipendenti dalla lingua e il loro valore semantico può essere compreso naturalmente e quindi consentono di riconciliare facilmente in un’unica ontologia i tipi di azione derivati dall’annotazione di corpora di lingue differenti (in particolare, in IMAGACT, di italiano e di inglese). Su questa base l’ontologia IMAGACT fornisce attraverso il riferimento a prototipi un quadro chiaro del *range* di variazione di un verbo azionale e un differenziale tra le azioni riferite da coppie di verbi senza operare definizioni né del senso dei verbi né dei tipi azionali.

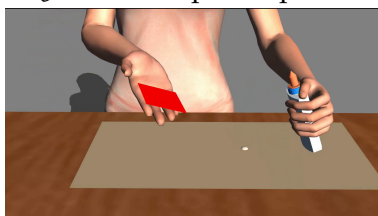
Al termine della fase di annotazione i tipi azionali sono collegati a delle scene (sotto forma di video) che determinano l’azione come prototipi rappresentativi del concetto azionale. Successivamente le scene sono trasformate in scene stereotipiche eliminando tutti i dettagli non essenziali alla rappresentazione del concetto. La forma utilizzata è il video 3D. La scena è l’immagine dell’azione fisica scelta per rappresentare prototipicamente un concetto azionale ed è pertanto un oggetto indipendente dalla lingua. Nel processo di mapping dell’annotazione su una ontologia di scene ognuna di esse può essere prototipo di tipi diversi e due tipi che categorizzano le stesse azioni si legano alla

stessa scena. Il rapporto tra l'espressione linguistica (il verbo) ed il suo riferimento non è però sempre così facile da esprimere attraverso un prototipo; sono in particolare due i fattori che aumentano la complessità:

- la diversa categorizzazione dello spazio azionale nelle lingue;
- il diverso livello di generalità dei verbi.

Il fatto che una stessa scena possa essere un prototipo per due tipi diversi non implica che questi due tipi debbano rappresentare lo stesso concetto azionale nello stesso modo. Nel caso in cui ci siano due tipi, uno più generale ed uno più specifico, la stessa scena può fungere infatti da prototipo per entrambi, ma i concetti che la scena deve rappresentare sono diversi: uno più esteso e uno più stretto. Quindi l'interpretazione della scena varia tra Tipo e Tipo: dal punto di vista del Tipo più generale il prototipo rappresenta una classe di azioni più ampia rispetto a quella che lo stesso prototipo rappresenta se lo si guarda dal punto di vista del verbo più specifico. Ad esempio la figura seguente (Fig. 15) mostra un'istantanea della scena "fe9fbeo9", che può fungere da prototipo sia per il verbo più generale *attaccare* (*Marta attacca il cartoncino sul foglio*), sia per il verbo più specifico *incollare* (*Marta incolla il cartoncino sul foglio*).

Figura 15: La scena prototipo di *incollare*



La stessa scena rappresenta un buon prototipo azionale per entrambi i verbi, che risultano localmente equivalenti all'interno di questa classe di azioni; i due tipi di *attaccare* e *incollare* non sono però equivalenti, ma bensì uno più generale e uno più specifico. Infatti un'azione come quella rappresentata in figura 16 può essere un altro prototipo dello stesso Tipo di *attaccare* (*Marco attacca il cartello al muro*), ma non dello stesso Tipo di *incollare*: nella scena non si incolla il cartello al muro, ma lo si inchioda.

Notiamo che dal punto di vista di *attaccare* non è significativo distinguere in tipi diversi sulla base dello strumento utilizzato per compiere l'azione (chiodi, colla, scotch,...),

Figura 16: La scena prototipo di *inchiodare*

mentre dal punto di vista di *incollare* è necessario. Questo non è un problema se si conosce il significato dei verbi, ma lo diventa quando si usano i prototipi azionali come mezzo per cercare una traduzione in una lingua non nota: in quel caso infatti possiamo essere sicuri dei verbi che possono predicare l'azione rappresentata nella scena prototipo, ma non possiamo sapere fino a che punto è possibile estendere il prototipo ad altre azioni simili, perché quei verbi continuano ad essere applicabili.

Sempre in riferimento all'esempio precedente, un apprendente dell'italiano, vedendo la figura 15 e la lista degli equivalenti, saprà con certezza di poter applicare sia *attaccare* che *incollare* a quell'azione specifica (IMAGACT offre questa informazione); ciò che invece non può sapere dal solo prototipo è l'ampiezza della classe di azioni che la scena rappresenta: in particolare non potrà giudicare se un'azione come quella di figura 16 fa parte dell'estensione dei due verbi oppure no.

Un altro problema, esposto più dettagliatamente nel capitolo 5, è quello della diversità nel modo di categorizzare lo spazio azionale: due tipi di due verbi di lingue diverse che si riferiscono alla stessa scena non ci danno informazioni riguardo ad eventuali relazioni semantiche che sussistono tra essi. È infatti possibile che i due tipi facciano riferimento a concetti azionali diversi e indipendenti, ma che trovino entrambi un buon rappresentante nel prototipo azionale. Un esempio di questo lo si trova nel capitolo 5, che propone un confronto tra il verbo cinese 盖 (*gài*) e il verbo italiano *abbassare*.

3.3 SIMILARITÀ PRAGMATICA E SEMANTICA TRA AZIONI: PROTOTIPI E FAMIGLIE DI PROTOTIPI

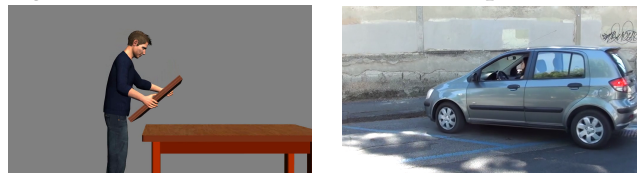
Una delle caratteristiche di IMAGACT è quella di fare ricorso alla capacità umana di giudicare la similarità. Il giudizio di similarità tra azioni - descritte da frasi non ambigue

– è il criterio con cui gli annotatori identificano i tipi azionali di un verbo. Anche a chi utilizza IMAGACT per scoprire come si traduce un verbo in un'altra lingua viene chiesto di scegliere il prototipo azionale più vicino all'azione che vuole esprimere con il verbo, quindi di fatto un giudizio di similarità tra l'azione da tradurre e le azioni esplicitate dalle scene. Di fatto lo scenario alla Wittgenstein su cui IMAGACT si fonda pone il giudizio di similarità al centro dello sfruttamento della risorsa.

La similarità tra azioni di IMAGACT è di due tipi, semantica e pragmatica e il confine tra queste ha margini di indeterminazione che debbono essere evidenziati. Un giudizio di similarità pragmatica è normalmente quello richiesto all'utilizzatore di IMAGACT per trovare l'azione più vicina a quella che questi vuole tradurre; solitamente per tipi non molto generali questo giudizio non presenta difficoltà: le classi di azioni ristrette a cui il verbo si estende comprendono azioni pragmaticamente simili e il traduttore troverà facilmente l'azione che sta cercando.

Viceversa i tipi definiti sulla base di proprietà semantiche più astratte avranno una variazione orizzontale è più variegata, e la similarità pragmatica tra azioni potrà essere scarsa. La similarità semantica delle azioni, cioè la loro aderenza a un concetto comune da un punto di vista pragmatico, può comprendere attività anche molto distanti. Le immagini seguenti, ad esempio, mostrano due azioni che fanno parte di uno stesso Tipo azionale di *mettere*, che è stato concepito sulla base di un tratto semantico molto generale, assolutamente necessario per la discriminazione di una classe ampia di variazioni di questo verbo e che può essere descritto come "collocazione di un oggetto in un luogo". Questa scelta fa sì che ci siano azioni pragmaticamente molto diverse come "parcheggiare un'auto" e "posare un quadro sul tavolo" che rientrano all'interno dello stesso concetto semantico.

Figura 17: Due scene di uno stesso Tipo di *mettere*



Questi casi pongono un problema relativamente allo sfruttamento del dizionario per

immagini di IMAGACT: infatti non è sempre facile reperire l'azione cercata, in particolare quando non si trova una scena pragmaticamente vicina nell'inventario di scene collegate a un verbo.

Un caso particolarmente critico è rappresentato dai cosiddetti *verbi logici*, quelli cioè la cui semantica è basata su una operazione logica tra gli argomenti (es. *aggiungere*, *unire*). In questi predicati la similarità semantica è l'unico criterio per l'identificazione dei tipi azionali: l'aspetto pragmatico non è rilevante e il traduttore deve fare un'induzione di senso per identificare il Tipo corretto. Il verbo *aggiungere*, ad esempio si fonda su un concetto puramente logico che può essere applicato a una varietà amplissima di azioni diverse che non siamo in grado di classificare in Tipi azionali sulla base della variazione del verbo. Le seguenti frasi sono estratte dal corpus di IMAGACT e, leggendole, ci si rende conto che non siamo in grado di classificarle sulla base di concetti pragmatici. In altre parole il criterio di similarità che consente la creazione di Tipi pragmaticamente coerenti non risulta applicabile e tali frasi sono state raggruppate in un unico insieme per mezzo di un tratto semantico astratto.

Fabio aggiunge l'acqua all'impasto

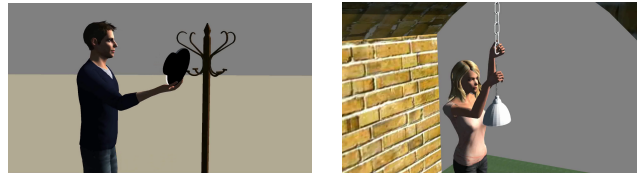
Il commerciante aggiunge un pacco al carico

Il cuoco aggiunge l'alloro al pollo arrosto

Sara aggiunge un mattoncino alla torre delle costruzioni

Come dimostrato dall'analisi di Tversky (1977) un giudizio di similarità è dipendente dal contesto, cioè dagli oggetti coinvolti: gli stessi due oggetti possono essere giudicati simili o dissimili tra loro se cambia il contesto di riferimento. Questo è stato spesso verificato durante la creazione dei tipi azionali in IMAGACT, in cui le stesse scene sono state giudicate simili rispetto a un verbo e dissimili rispetto ad un altro verbo. Ad esempio dal punto di vista del verbo *appendere*, le due scene di Figura 18 sono prototipi di due Tipi diversi, poiché viene considerato differenziale il tratto di "mettere in sospensione". In altri termini osservando come varia il verbo *appendere* c'è interesse a distinguere i tipi di eventi in cui *appendere* è equivalente a *sospendere*, dagli eventi in cui non lo è.

Questa distinzione non è però significativa per il verbo più generale *attaccare*, rispetto al quale queste due azioni fanno parte della variazione orizzontale e quindi sono raggrup-

Figura 18: Due scene di *appendere*

pate all'interno di un unico Tipo azionale. Infatti, nel caso della variazione di *attaccare* c'è interesse a distinguere primariamente, ad esempio, quando l'azione di *attaccare* è equivalente ad *incollare* rispetto a quando l'azione di *attaccare* è equivalente ad *appendere* e la distinzione precedente risulta quindi troppo granulare.

In questo processo la competenza semantica dell'annotatore madrelingua è servita da base per i giudizi di similarità: ciò che l'annotatore ha fatto durante la creazione di tipi è stata un'operazione di *clustering* di azioni nel contesto della semantica verbale. Il task che è stato seguito per la creazione dei tipi azionali può essere esplicitato come segue:

Dato un verbo v e due azioni a_1 e a_2 predicabili con v , ciascuna esplicitata da una frase non ambigua, l'annotatore giudica se "dal punto di vista del verbo v , le azioni a_1 e a_2 sono simili".

Dove il testo tra virgolette rappresenta la funzione di *clustering* che discrimina se due azioni fanno parte o meno dello stesso Tipo. Qui "dal punto di vista del verbo v " è la base contestuale a cui l'annotatore, in quanto madrelingua competente, ha accesso e che rafforza il giudizio di similarità.

Allo stesso modo l'apprendente che sfrutta l'ontologia per immagini per tradurre un verbo in un'altra lingua si trova a dover giudicare la similarità dell'azione che vuole descrivere con quella rappresentata nei prototipi collegati al verbo. Questo è il task che viene eseguito per ogni scena prototipo collegata al verbo in lingua *source*:

Data un'azione a_1 per cui si cerca un verbo in lingua *target* e un'azione a_2 , descritta dal prototipo azionale, l'apprendente giudica "se a_1 e a_2 sono simili".

Come si vede, qui non è presente nessuna informazione lessicale che può guidare il giudizio di similarità (e che nel caso precedente era rappresentata dalla presenza del verbo *appendere*, rispetto al quale il parlante ha una competenza lessicale). Pertanto

il giudizio di similarità non sarà sempre efficace: in particolare esso sarà adeguato per i tipi più specifici, cioè quelli in cui l'estensione è ristretta e circoscritta ad azioni pragmaticamente molto simili, ma inadatto per i tipi più generali.

Questa difficoltà è una delle motivazioni che stanno alla base della modificazione del modello generale di rappresentazione dell'ontologia (insiemistico) proposto in questa tesi di dottorato. Si deve notare infatti che, se al posto di un solo prototipo si utilizza un insieme di scene prototipiche (famiglia di prototipi), il task viene modificato come segue:

Data un'azione a_1 per cui si cerca un verbo in lingua *target* e una serie di azioni a_2, \dots, a_n , descritte da prototipi azionali, l'apprendente giudica "se a_1 è simile a a_2, \dots, a_n ".

In questa nuova formulazione l'informazione che a_2, \dots, a_n siano tra loro simili rappresenta la base cognitiva su cui l'utente si trova a giudicare la similarità. In altre parole la domanda di giudizio diviene: "dal momento che queste azioni a_2, \dots, a_n sono tutte simili tra loro, si può dire che anche a_1 fa parte del gruppo?". In tal modo, se l'insieme dei prototipi (o famiglia) è sufficientemente rappresentativo, l'utente può determinare in quale gruppo di scene si andrebbe ad inserire l'azione che vuole rappresentare.

Tversky (1977) mostra come l'aggiunta di nuovi oggetti all'interno di una categoria causi la rimozione di alcune *features* (quelle non condivise da tutti gli elementi) giudicate rilevanti per la categorizzazione. In IMAGACT questa proprietà è sfruttata per eliminare quei tratti della scena che sono irrilevanti per il Tipo e, parallelamente, per aumentare la rilevanza delle *features* comuni condivise da tutte le scene appartenenti all'insieme. Così nei tipi più generali troveremo una serie di scene prototipo che possono essere anche molto diverse tra loro, ma che hanno in comune alcune *features* molto generali.

Grazie all'utilizzo di "insiemi di prototipi", riuniti in famiglie IMAGACT riesce quindi ad esplicitare l'estensione del Tipo generale, che risulta facilmente giudicabile per similarità pragmatica e semantica. In questi i casi infatti i due tipi di similarità vengono utilizzati: l'insieme di scene mostrate consentono all'utilizzatore di indurre il concetto generale identificato dal Tipo e quindi di poter giudicare se l'azione cercata appartiene o meno ad esso; parallelamente le diverse scene identificano i nuclei pragmatici prominenti della variazione orizzontale e quindi potrà essere trovata una scena simile (pragmaticamente) a quella cercata.

Quanti e quali prototipi utilizzare per rappresentare un Tipo azionale non è facilmente determinabile; IMAGACT utilizza le gerarchie di tipi per poter indurre questa informazione, considerando tutte le scene prototipo sotto-ordinate (prototipi di tipi più specifici) come prototipi del Tipo generale. Questo approccio fa in modo che il numero di scene prototipiche sia dipendente dalla generalità del Tipo e si basa sul seguente assunto:

I nuclei della variazione orizzontale di un verbo possono essere riferiti da verbi più specifici nella stessa lingua

Questo assunto è stato verificato in molti casi, proprio per il forte legame che sussiste tra la lingua e la categorizzazione dello spazio azionale. Uno dei motivi per cui siamo in grado di distinguere un concetto azionale indipendente all'interno dell'estensione di un verbo è proprio perché abbiamo un verbo specifico per descriverlo, ma anche, dal punto di vista opposto, diremmo che la necessità pragmatica di riferirsi ad un nucleo pragmatico specifico determina l'esistenza di un verbo in grado di categorizzarlo.

La strategia che utilizza l'interfaccia di IMAGACT per la rappresentazione del riferimento è quella di individuare i tipi più specifici attraverso un prototipo e quelli più generali con una serie di scene prototipiche, una per ogni nucleo pragmatico della variazione orizzontale.

Può essere interessante segnalare in questa discussione che questo metodo di rappresentazione non è efficace per i verbi di azioni logiche (*unire, aggiungere* ecc.), per i quali la similarità pragmatica tra azioni è irrilevante. Con questi tipi un qualsiasi numero di scene sarebbe insufficiente a fornirne una rappresentazione adeguata perché potrà sempre essere trovata un'azione appartenente al Tipo, ma dissimile da tutte le altre: con questi verbi non possiamo evidenziare i nuclei pragmatici perché non ci sono. I verbi logici, in altri termini, codificano concetti azionali non esprimibili con immagini⁵.

⁵ In IMAGACT sono presenti alcuni verbi logici, come *unire* o *aggiungere*; essi hanno un solo Tipo azionale contenente alcune scene, che in questo caso non hanno il ruolo di prototipi, ma di esempi. Per questi verbi la rappresentazione tramite serie di immagini funziona soltanto se nella lingua dell'utente è lessicalizzato un concetto simile a quello che sta dietro al verbo logico; in tal caso l'utente, non trovando similarità pragmatica tra le scene, ma avendo nella sua competenza semantica un concetto in grado di categorizzare quella serie di azioni, sarà indotto a legare al Tipo quel concetto. In tal modo potrà quindi capire il concetto azionale e giudicarlo la similarità semantica con l'azione che intende tradurre.

3.4 LE INFERENZE

Una conseguenza importante del modo di rappresentare le relazioni tra verbi attraverso prototipi e famiglie sono le modalità con cui possono essere espresse le relazioni inferenziali tra verbi. Infatti, ciò che rende, in teoria, l'ontologia IMAGACT una risorsa sfruttabile per la disambiguazione e la traduzione dei verbi generali è la possibilità di utilizzare il riferimento azionale per formulare inferenze logiche. Come vedremo (Cap. 4) il potere inferenziale di IMAGACT deriva proprio dalla suddivisione in tipi azionali. In questo quadro teorico l'inferenza si sviluppa attraverso il principio di equivalenza locale:

- (1) se due tipi si riferiscono alla stessa classe di azioni allora i verbi ad essi collegati sono localmente equivalenti

Oltre al caso di tipi equivalenti ci sono anche rapporti di Tipo gerarchico tra tipi, che possono essere tra loro legati dal più generale al più specifico; in questi casi il verbo collegato al Tipo più generale è applicabile all'intera classe di azioni del Tipo più specifico, ma non è vero il viceversa.

- (2) se un Tipo t_1 è più generale di un Tipo t_2 allora i verbi collegati a t_1 sono localmente equivalenti ai verbi collegati a t_2

Notiamo quindi che dove c'è equivalenza locale tra due predicati all'interno di un contesto pragmatico è possibile formulare inferenze del Tipo $A \Leftrightarrow B$ (nel caso 1) oppure $A \Rightarrow B$ (in 2).

Per fornire una descrizione non ambigua di queste formulazioni logiche dobbiamo chiarificare cosa siano A e B , questione non banale dal momento che la concezione delle entità di base di IMAGACT, in particolare i tipi azionali e le scene prototipiche, è molto cambiata nello sviluppo dell'ontologia. Senza entrare nel dettaglio di questioni strutturali che saranno trattate nei capitoli successivi, diciamo che ciò che costituisce il fondamento di IMAGACT, cioè la determinazione dei concetti azionali su base pragmatica e il principio di equivalenza locale, ci consente di dire che le inferenze si sviluppano tra Tipo e verbo e questa è una proprietà generale e invariante dell'ontologia. Pertanto, dato un Tipo t di un verbo v_1 e un verbo v_2 localmente equivalente a v_1 nel contesto pragmatico rappresentato da t , possiamo sempre dire che "se l'azione appartiene al Tipo

t allora è possibile applicare v_2 per predicarla" e questo vale indipendentemente dal modello scelto per la rappresentazione dell'informazione.

Su questa regola generale poi i diversi modelli teorici sviluppati sopra IMAGACT definiscono delle regole più specifiche in termini di rapporti tra scene e tipi e ciascuno di essi assegna un valore diverso agli elementi A e B delle espressioni (1) e (2). Le formalizzazioni delle inferenze nel modello gerarchico originale e nella nuova struttura dati insiemistica sono riportate rispettivamente in 4.3.5 e in 6.1.2.

IL MODELLO GERARCHICO DI IMAGACT

4.1 ENTITÀ E RELAZIONI PRINCIPALI

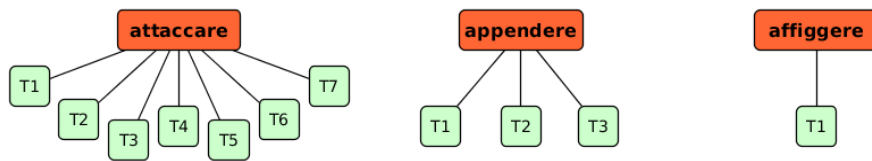
IMAGACT contiene una ampia varietà di informazioni, dai segmenti testuali provenienti da corpus a tutti i dati linguistici derivanti dai diversi livelli di annotazione. Tutta questa ricchezza informativa è legata ad una struttura dati centrale, che si basa su un modello logico gerarchico. Il cuore dell'ontologia è rappresentato da tre entità principali: il Verbo, il Tipo e la Scena. Questi oggetti e le relazioni che li collegano tra loro formalizzano il rapporto tra il linguaggio e l'azione fisica e costituiscono quindi il nucleo del modello teorico su cui si fonda IMAGACT.

IL VERBO Il Verbo è propriamente un lemma verbale di una lingua: sotto questa voce rientrano sia i verbi, sia le costruzioni verbali che hanno indipendenza semantica, come ad esempio i *phrasal verb* inglesi o i verbi sintagmatici dell'italiano. Le forme riflessive dell'italiano e, in generale, tutte quelle modificazioni del verbo che non portano ad un'alterazione del significato vengono riportate al lemma principale, come avviene normalmente nei dizionari. In realtà esiste un problema nell'identificare cosa sia un lemma verbale nelle diverse lingue, ma non è oggetto di questa tesi (si veda per la lingua cinese Pan, res).

IL TIPO Come abbiamo visto ogni verbo può avere uno o più tipi azionali che ne definiscono gli ambiti di variazione verticale. I verbi con un solo Tipo sono privi di variazione verticale, mentre un numero elevato di tipi evidenzia una alta generalità e, parallelamente, una forte ambiguità. Ogni Tipo rappresenta un concetto azionale all'interno della variazione di un verbo e identifica pertanto una classe di azioni che, normalmente, è pragmaticamente coesa e ben identificabile in qualsiasi lingua. Entrambe

queste due caratteristiche non sono valide in assoluto: come abbiamo visto (Par. 3.3) alcuni verbi hanno tipi con scarsa coerenza pragmatica e, nel par. 5.2 sono mostrati casi in cui un concetto azionale chiaro per un parlante italiano non è condiviso in cinese. In generale però la discriminazione dei tipi azionali fornisce una base importante di disambiguazione in contesto multilingue.

Figura 19: La suddivisione in Tipi azionali dei lemmi verbali

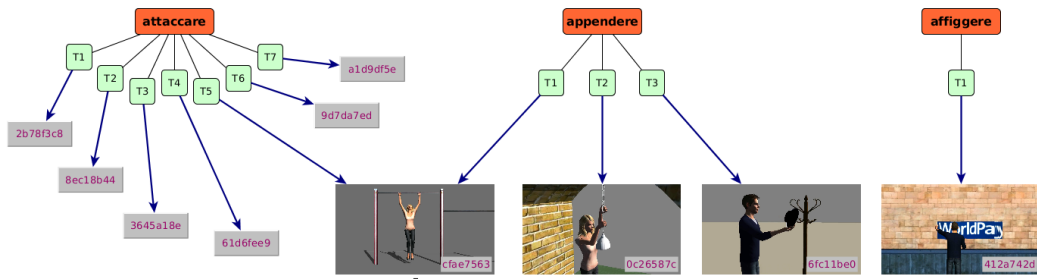


LA SCENA Una Scena è un'istanza prototipica della classe di azioni a cui un Tipo si riferisce: il formato utilizzato è il video 3D, che rappresenta una azione, considerata un buon prototipo per elicitarne il concetto identificato dal Tipo. IMAGACT attualmente contiene 1.010 scene incluse le scene vuote (vedi par. 4.3.3).

LE RELAZIONI DI PROTOTIPO E ISTANZA Ogni Tipo è collegato ad una e una sola Scena tramite la relazione di *prototipo*; questa relazione esprime il fatto che la Scena sia un prototipo rappresentativo della classe di azioni definita dal Tipo.

Mentre il prototipo è una relazione *1-a-1* tra il Tipo e la Scena, la relazione inversa è *1-a-molti*; una Scena può infatti essere un prototipo di tipi diversi (di verbi diversi).

Figura 20: La relazione di *prototipo* tra Tipi e Scene



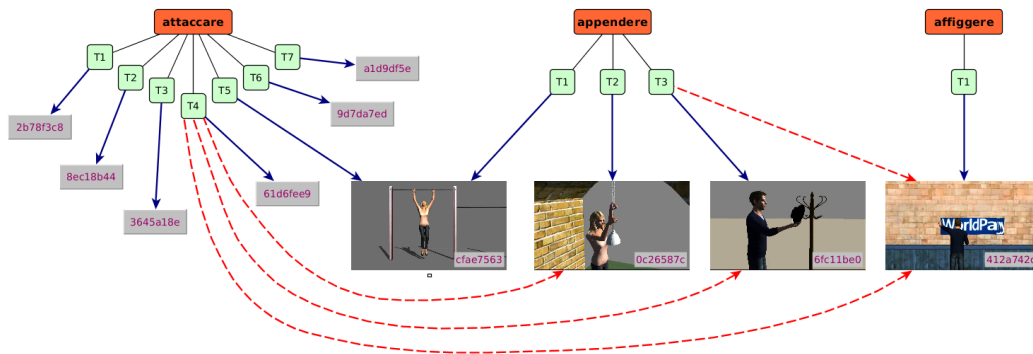
In questo modello l'equivalenza di due tipi è rappresentata dalla condivisione di un prototipo: quindi due tipi collegati con la relazione di *prototipo* alla stessa Scena

categorizzano le stesse azioni azioni. Di conseguenza i verbi collegati sono *localmente equivalenti*.

Nell'esempio di figura 20 i tipi *appendere-T1* e *attaccare-T5* sono equivalenti: essi infatti rappresentano la stessa classe di azioni e condividono pertanto lo stesso prototipo azionale, rappresentato dalla Scena "cfae7563". Ne deriva che i verbi *attaccare* e *appendere* siano localmente equivalenti in quel luogo.

Oltre al prototipo esiste un'altra relazione che collega i tipi alle scene: la relazione di *istanza*. Questa è *1-a-molti* e determina il fatto che la classe di azioni rappresentata dalla Scena sia una sottoclasse delle azioni categorizzate da un Tipo azionale: se una Scena è *istanza* di un Tipo, essa rappresenta un prototipo di un sottoinsieme degli eventi categorizzati dal Tipo. Anche la relazione di *istanza* tra Scena e Tipo implica l'applicabilità linguistica del verbo di cui il Tipo fa parte alla classe di azioni rappresentata dalla Scena. Ogni verbo è quindi applicabile nel suo senso proprio a tutte le azioni rappresentate dalle scene ad esso collegate come prototipo o come istanza ai suoi tipi.

Figura 21: La relazione di *istanza* tra Tipi e Scene



Se si aggiungono le relazioni di istanza all'esempio precedente (Fig. 21) vediamo che le scene "412a742d", "6fc11be0" e "0c26587c", oltre ad essere prototipi rispettivamente di un Tipo di *affiggere* e di due tipi di *appendere*, sono anche possibili istanze di *attaccare-T4*. Inoltre "412a742d" è anche istanza di *appendere-T3*. Poiché anche la relazione di *istanza* determina l'equivalenza locale, allora abbiamo la seguente situazione:

- *affiggere*, *appendere* e *attaccare* sono localmente equivalenti nel campo di variazione di "412a742d";

4.2 DERIVAZIONE DEL MODELLO GERARCHICO

- *appendere* e *attaccare* sono localmente equivalenti nei campi di variazione di “cfae7563”, “oc26587c” e “6fc11be0”.

In questo caso particolare possiamo formulare anche inferenze dirette tra verbi:

se affiggo allora appendo

se affiggo allora attacco

se appendo allora attacco

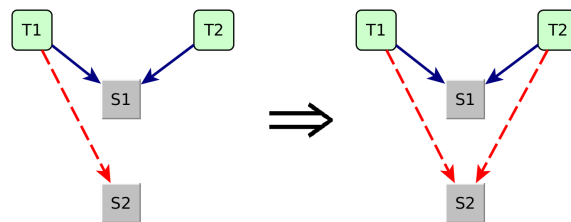
Notiamo infatti che tutti i tipi di *affiggere* sono istanze di scene prototipo di *appendere* e tutti i tipi di *appendere* sono collegati ad un prototipo di *attaccare* o come istanza (*appendere-T2* e *appendere-T3*) o come prototipo (*appendere-T1*).

4.2 DERIVAZIONE DEL MODELLO GERARCHICO

Durante la fase di annotazione sono stati imposti dei vincoli alla relazione di *istanza*, che hanno permesso di formalizzarne l'utilizzo:

1. Se una Scena s_1 è prototipo dei due tipi t_1 e t_2 , e il Tipo t_1 ha anche una Scena s_2 collegata come istanza, allora s_2 è istanza anche di t_2 e in generale di tutti i tipi che condividono lo stesso prototipo.

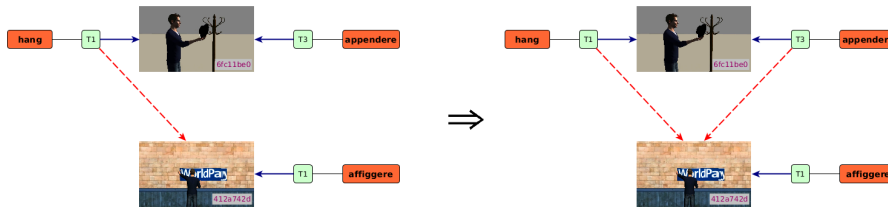
Figura 22: Prima regola sull'applicazione della relazione di *istanza*.



Ad esempio la Scena “412a742d”, che costituisce il prototipo del Tipo 1 di *affiggere*, è anche un’*istanza* del Tipo 1 del verbo inglese *hang*. Per la regola sopracitata la stessa Scena è *istanza* anche del Tipo 3 di *appendere*, dal momento che esso ha lo stesso prototipo del Tipo 1 di *hang*.

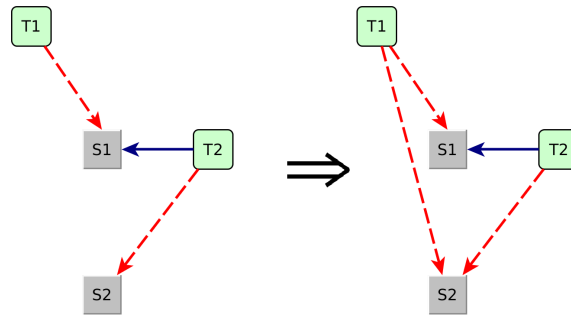
4.2 DERIVAZIONE DEL MODELLO GERARCHICO

Figura 23: Esempio d'uso della regola sui verbi *appendere* e *affiggere*



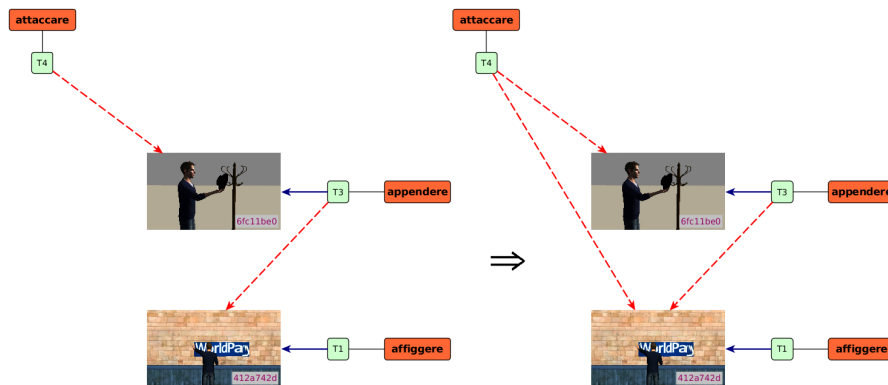
2. Se un Tipo t_1 ha come *istanza* una Scena s_1 , e la Scena s_1 è prototipo di t_2 , la Scena s_2 che è *istanza* di t_2 è anche *istanza* di t_1 .

Figura 24: Seconda regola sull'applicazione della relazione di *istanza*.



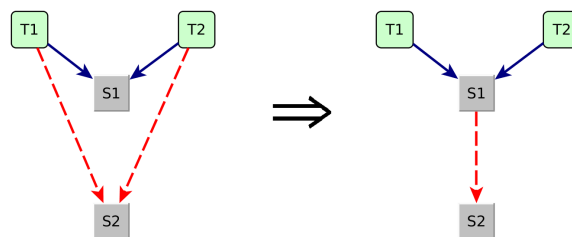
Nell'esempio seguente si vede che la Scena di *affiggere* viene ad essere *istanza* di un Tipo di *attaccare* in virtù della catena di relazioni di *istanza* che sussiste tra i Tipi di *attaccare*, *appendere* e *affiggere* in base alla regola (2).

Figura 25: Esempio d'uso della regola sui verbi *appendere* e *affiggere*



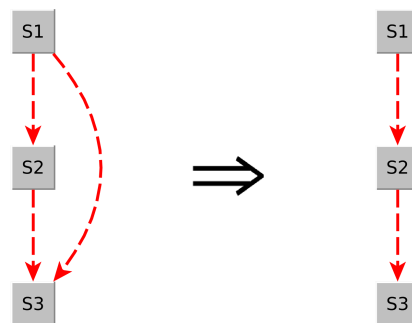
Queste due restrizioni conferiscono alla relazione di *istanza* un carattere transitivo: essa diventa propriamente transitiva quando la si considera come una relazione tra due Scene, anziché tra Tipo e Scena. La (1) permette infatti di spostare sull'oggetto Scena la parte destra della relazione senza alcuna perdita di informazione; in questo modo si introduce una prima semplificazione: tutti i collegamenti di *istanza* tra una Scena s e n Tipi localmente equivalenti vengono sostituiti con un'unica relazione tra la Scena s e la Scena prototipo degli n Tipi.

Figura 26: Istanza come relazione tra due Scene



La (2) consente di introdurre una seconda semplificazione, dal momento che determina la transitività della relazione di *istanza*; pertanto vengono eliminate tutte le *istanze* tra Scene che possono essere inferibili dalla transitività.

Figura 27: Istanza come relazione transitiva

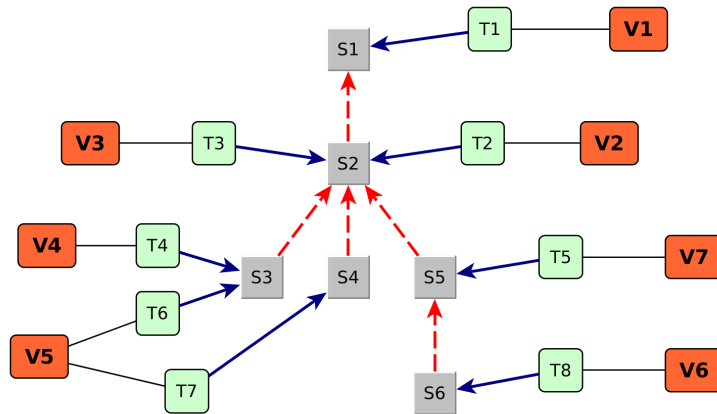


Il cambiamento di dominio della relazione di *istanza* è una trasformazione che produce un modello dati equivalente a quello derivante dall'annotazione, ma meno ridondante e più facilmente interpretabile, in quanto permette di avere un'immagine chiara della gerarchia implicita tra Scene che sta sotto IMAGACT. La struttura dati così generata è un albero di Scene tra loro collegate attraverso la relazione di *istanza* dalla più generale

4.3 PROPRIETÀ DEL MODELLO GERARCHICO

(la radice) alla più specifica (le foglie). Ad essa ogni Tipo si collega (con la relazione di *prototipo*) nel punto dell'albero che ne rappresenta correttamente il riferimento.

Figura 28: Il modello gerarchico di IMAGACT



La concettualizzazione derivata corrisponde al modello logico gerarchico di IMAGACT; di seguito si farà sempre riferimento ad essa e non alla rappresentazione fisica dei dati che vede l'*istanza* come una relazione tra Scena e Tipo.

4.3 PROPRIETÀ DEL MODELLO GERARCHICO

4.3.1 La Scena: concetto e rappresentazione

Come abbiamo visto, nel modello dati di IMAGACT la relazione di prototipo sussistente tra un Tipo e una Scena è strettamente legata al riferimento comune alle azioni: due tipi collegati con la relazione di prototipo alla stessa Scena categorizzano la stessa classe di azioni e viceversa due tipi che categorizzano la stessa classe di azioni sono entrambi collegati ad una Scena come prototipo. Questa assunzione conferisce implicitamente all'oggetto Scena un ruolo concettuale. Se di per sé infatti una Scena non è altro che la rappresentazione in forma di video di una specifica azione fisica, essa diviene il rappresentante di un (e un solo) concetto azionale specifico, quello cioè definito dai tipi che ad essa si collegano con la relazione di prototipo. Questo genera un certo grado di sovrapposizione tra il Tipo e la Scena; ne possiamo avere un'evidenza ancora più

immediata se consideriamo la relazione di istanza che, nella riformulazione della struttura dati esemplificata in Fig. 28, è una relazione transitiva tra Scene. Ciò significa che alcune scene sono più specifiche di altre, cosa che non ha nessun senso se consideriamo la Scena unicamente come la rappresentazione di un'azione fisica; ovviamente le relazioni gerarchiche che intercorrono tra gli oggetti Scena non si riferiscono strettamente alle azioni, ma ai concetti azionali, per i quali è perfettamente sensata una strutturazione tassonomica che tenga conto del loro diverso livello di generalità. È importante tenere in considerazione questo doppio ruolo che ha la Scena in IMAGACT: essa infatti da un lato è l'oggetto indipendente dalla lingua che fa da tramite per modellare i rapporti tra verbi in lingue diverse; dall'altro rappresenta uno specifico concetto azionale e quindi, poiché le lingue categorizzano lo spazio in modo diverso, è definita sulla base di una lingua e potrebbe essere una concettualizzazione inadeguata per un'altra lingua. Questa ambiguità si è sviluppata nella costruzione dell'ontologia e solo in pochi casi ha generato problemi di rappresentazione: in particolare il lavoro di mapping tra l'italiano e l'inglese ha mostrato che molti concetti sono condivisi tra le due lingue e, anche quando non è così, normalmente sono compatibili con la classificazione degli altri concetti e quindi possono essere incapsulati nel modello gerarchico. La questione è però molto diversa per lingue "lontane", in cui il ruolo ambiguo della Scena porta a delle contraddizioni e alla incapacità di fornire una rappresentazione coerente dei concetti (si veda in particolare il Par. 5.2 sui problemi di estensione alla lingua cinese).

4.3.2 *Equivalenza locale e famiglie di prototipi*

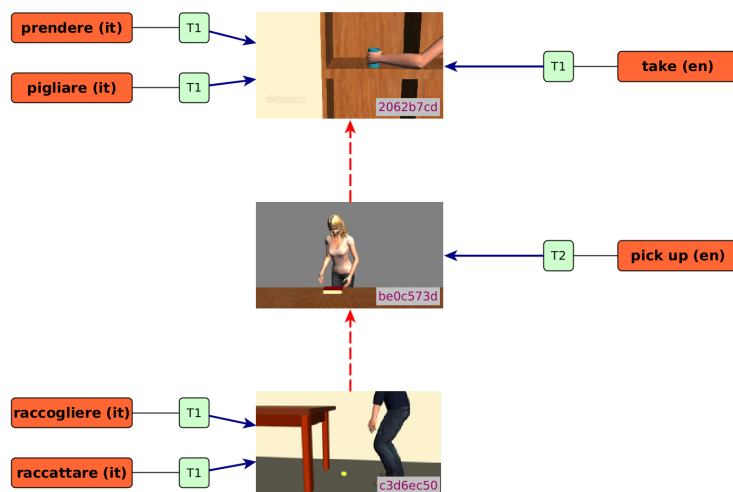
In questo modello dati, data una Scena S ,

- tutti i verbi collegati (attraverso i loro tipi) a S con la relazione di prototipo sono localmente equivalenti;
- tutti i verbi collegati (attraverso i loro tipi) a scene di cui S è un'istanza sono localmente equivalenti.

L'esempio seguente (Fig. 29) mostra le relazioni che intercorrono tra *raccogliere*, *prendere*, *take* e *pick up*. Un Tipo di *prendere*, un Tipo di *pigliare* e un Tipo di *take* fanno riferimento

ad uno stesso concetto azionale molto generale rappresentato dalla Scena “2062b7cd” (*Fabio prende la tazza dal ripiano*). Ad un livello inferiore troviamo un Tipo di *pick up*, che si può applicare soltanto se il tema si trova in una posizione più bassa rispetto all’agente, ed è rappresentato dalla Scena “be0c573d” (*Mary picks up the book from the table*). In fondo alla gerarchia ci sono i verbi italiani *raccogliere* e *raccattare*, che sono ancora più specifici, poiché richiedono che il tema si trovi per terra (Scena “c3d6ec50”, *Fabio raccoglie la pallina da terra*).

Figura 29: Relazioni di equivalenza locale tra *prendere*, *raccogliere* e *pick up*



Per le proprietà sopra descritte i verbi *prendere*, *pigliare*, *take*, *pick up*, *raccogliere* e *raccattare* sono tra loro localmente equivalenti: essi infatti possono tutti essere applicati per descrivere la Scena “c3d6ec50” e sono quindi equivalenti nel campo di variazione da essa rappresentato.

La Scena viene arricchita dall’alto con tutti i verbi collegati (attraverso uno dei loro tipi) alle scene sopra-ordinate

Ovvero, data una Scena, l’azione che essa rappresenta può essere predicata non soltanto dai verbi ad essa collegati direttamente (cioè con la relazione di *prototipo*), ma anche da tutti i verbi che si mappano sui livelli più alti della gerarchia.

Come abbiamo già accennato nel paragrafo 3.3, i concetti generali non sono rappresentabili in modo efficace attraverso un solo prototipo, che non riesce ad esprimere chiaramente tutta la variazione orizzontale. Per questa ragione IMAGACT utilizza i

prototipi di classi di azione più specifiche per arricchire la rappresentazione del concetto generale. In questa struttura i concetti più specifici sono identificati dalle scene sotto-ordinate, cioè quelle collegate attraverso la relazione di istanza. Esse infatti sono istanze azionali possibili del concetto più generale espresso dalla Scena al livello superiore. Per questa ragione in IMAGACT, la rappresentazione di un concetto generale è fornita dalla Scena prototipo unita a tutti i prototipi delle scene collegate ad essa con la relazione di istanza, ad un qualsiasi livello di profondità (si ricorda che la relazione di istanza è transitiva).

Nell'esempio di figura 29 la coppia formata dalle scene "c3d6ec50" e "be0c573d" costituisce la famiglia di prototipi per il Tipo di *pick up*, mentre l'insieme di tutte e tre le scene è la famiglia di prototipi per i tipi più generali di *prendere*, *pigliare* e *take*.

La Scena viene arricchita dal basso con tanti prototipi quanti sono le azioni rappresentate dalle scene sotto-ordinate

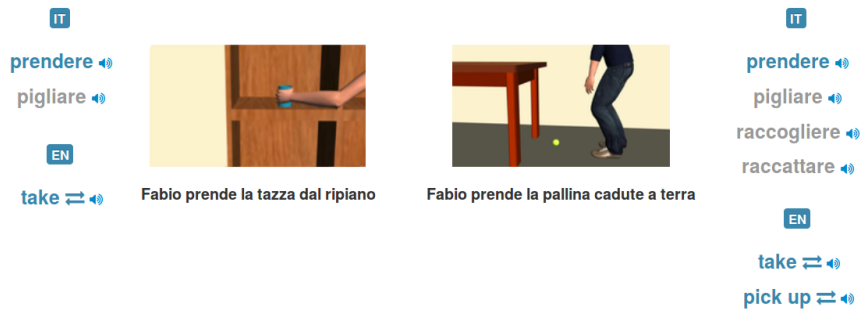
Ovvero, data una Scena, tutte le scene che si trovano ai livelli più bassi della gerarchia possono fungere da prototipo per il concetto azionale che essa rappresenta.

Vediamo quindi che c'è un rapporto inverso tra il numero di verbi applicabili ad un concetto azionale e il numero di prototipi che lo descrivono: i concetti molto generali infatti sono rappresentati da un alto numero di scene, ma hanno pochi verbi collegati; viceversa i concetti più specifici hanno una sola Scena prototipica e molti verbi per predicarla. La figura seguente (Fig. 30) riassume questa condizione: il prototipo a destra rappresenta un concetto più specifico che è predicabile con molti verbi, ma identificabile da un solo prototipo. Viceversa la Scena a destra ha pochi verbi collegati, poiché rappresenta un concetto generale; tutte e tre le scene di figura 29 possono però alternativamente fungere da prototipi.

4.3.3 Scene vuote

Le cosiddette *scene vuote* sono oggetti Scena privi di immagine: esse sono un altro esempio del ruolo concettuale conferito alla Scena in questa struttura dati. Dal momento che questi oggetti non contengono il riempimento azionale, essi rappresentano unicamente

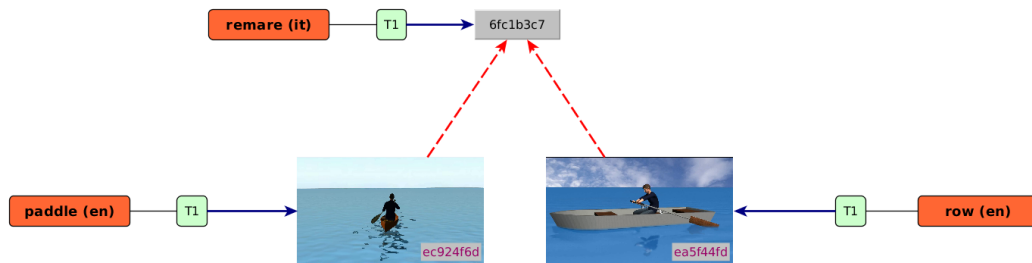
Figura 30: Concetti generali e specifici



un concetto; il riferimento all'azione in questi casi è rappresentato dall'insieme delle scene "piene" che si trovano ai livelli inferiori della gerarchia.

Le scene vuote sono state inserite in tutti quei casi in cui le scene sotto-ordinate fornivano una rappresentazione completa della variazione orizzontale di un Tipo. In questi casi l'aggiunta di un nuovo prototipo alla famiglia sarebbe ridondante, poiché i nuclei della variazione orizzontale sono già rappresentati.

Figura 31: La Scena vuota prototipo di *remare*



Nell'esempio di figura 31 si vede come i prototipi azionali di due tipi che fanno riferimento ai verbi inglesi *row* e *paddle* forniscano una rappresentazione sufficientemente chiara del significato di *remare* (che è più generale e ha un solo Tipo). Pertanto la Scena "6fc1b3c7", che eredita i prototipi sotto-ordinati, non necessita di un riempimento azionale ed è stata lasciata vuota. Nel database di IMAGACT l'utilizzo di scene vuote è diffuso specialmente per i concetti azionali molto generali in cui sono presenti più livelli sotto-ordinati; solitamente in questi casi l'insieme delle scene che costituiscono la famiglia di prototipi è sufficientemente ricco e consente di fare a meno di un ulteriore prototipo. Quando invece la famiglia di prototipi sotto-ordinati non è sufficiente a coprire

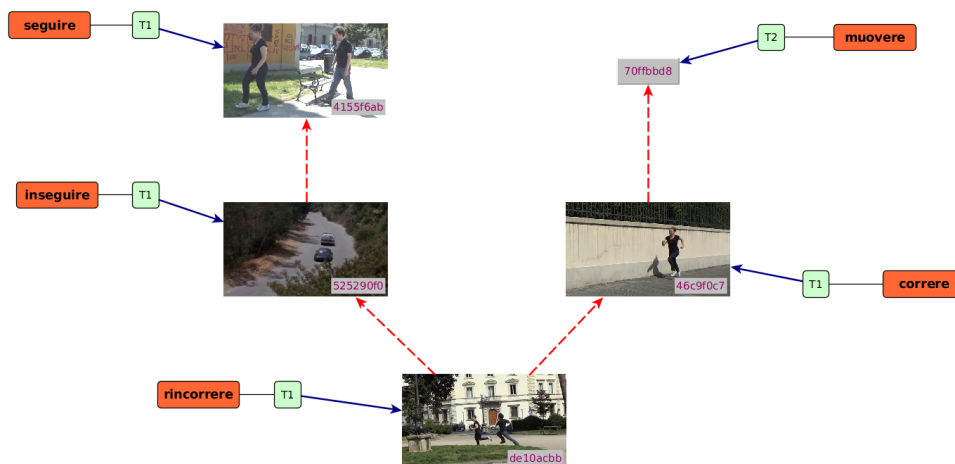
tutta la classe di azioni del Tipo azionale, allora la Scena viene riempita con un ulteriore prototipo.

4.3.4 Scene figlie di due scene

Sebbene il modello dati sopra descritto sia quasi sempre rappresentabile in una struttura ad albero, non ci sono vincoli sul numero di scene sopra-ordinate. Questo caso si presenta per 12 scene di IMAGACT, ciascuna delle quali è istanza di 2 scene anziché di una sola.

Il seguente esempio mostra il caso di *rincorrere* (un solo Tipo), in cui si ha una relazione di istanza con due concetti azionali separati: quello di *inseguire* e quello di *correre*. In altri termini si vuol cogliere il fatto che *rincorrere* è sia *inseguire* che *correre*.

Figura 32: Le due scene sopra-ordinate di *rincorrere*



Questi casi non presentano un problema e tutte le proprietà del modello continuano a valere:

- *seguire*, *inseguire*, *muovere*, *correre* e *rincorrere* sono tra loro localmente equivalenti, poiché sono tutti applicabili alla Scena "de10acbb";
- il Tipo 1 di *seguire* è prototipicamente rappresentato dalla famiglia formata dalle tre scene "de10acbb", "525290fo" e "4155f6ab"; le due scene "46c9f0c7" e "de10acbb" sono entrambi prototipi del Tipo 1 di *correre*.

4.3.5 *Le inferenze*

A partire da quanto accennato nel Par. 3.4 vediamo adesso come la struttura dati sia in grado di permettere le inferenze. Le relazioni tra verbi, tipi e scene consentono di formulare delle implicazioni logiche tra verbi sia intra- che inter-linguistiche come le seguenti:

(1) *se aggancio allora attacco*

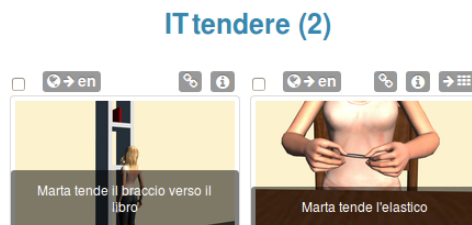
(2) *se aggancio allora hook*

In (1) l'inferenza è intra-linguistica ed è derivata dalla relazione di istanza tra la Scena prototipo dell'unico Tipo di *agganciare* (più specifica) e la Scena prototipo di un Tipo di *attaccare* (più generale). In (2) l'implicazione è tra un verbo italiano ed uno inglese e sussiste poiché il Tipo di *agganciare* e un Tipo di *hook* condividono lo stesso prototipo: sono cioè localmente equivalenti.

Inferenze dirette tra verbi come questa sono un caso particolare ed abbastanza raro, poiché possono essere fatte soltanto quando un verbo è applicabile all'intera variazione verticale di un altro verbo. L'enunciato (3), ad esempio, risulta essere falso, poiché, benché sussista una relazione di istanza tra il prototipo di un Tipo di *tendere* e il prototipo di un Tipo di *tirare*, il verbo *tendere* ha due tipi azionali (Fig. 33) ed uno di essi non è collegato a *tirare*.

(3) *se tendo allora tiro*

Figura 33: I due Tipi di *tendere* nell'interfaccia web di IMAGACT



Nello specifico la Scena di "Marta tende l'elastico" è prototipo di un Tipo di *tendere* (Tipo 1) e istanza di un Tipo di *tirare*, mentre la Scena di "Marta tende un braccio verso il libro" è prototipo di un altro Tipo di *tendere* (Tipo 2), ma non ha relazioni con alcun Tipo

di *tirare*. Questo rende impossibile avere un'implicazione diretta tra verbi, ma permette di inferire che ogni volta che *tendere* è usato per predicare azioni dello stesso Tipo di "Marta tende l'elastico", allora è possibile usare anche *tirare*. La (3) può quindi essere resa vera se riformulata con l'informazione relativa al Tipo:

(4) *se tendo-T₂ allora tiro*

Allo stesso modo è possibile realizzare l'inferenza negativa:

(5) *se tendo-T₁ allora non tiro*

Grazie alla presenza dei tipi è quindi possibile formulare una quantità di implicazioni logiche tra verbi che altrimenti non sarebbero possibili, a causa dell'ambiguità di significato dei verbi nelle lingue. Il concetto stesso di inferenza è rimodulato in IMAGACT su base pragmatica: l'ontologia non rappresenta infatti relazioni logiche tra i sensi dei verbi, ma tra i tipi, che sono una partizione del senso realizzata su base pragmatica. Questo incrementa in modo notevole il potere inferenziale dell'ontologia, che permette di esprimere relazioni, come ad esempio (4) e (5), che ci appaiono perfettamente sensate, ma che non possono essere formalizzate senza una suddivisione tipologica del significato dei verbi.

Questi casi possono essere più formalmente descritti come segue.

- **Inferenza diretta tra verbi**

Dati due verbi v_1 e v_2 , l'enunciato "se v_1 allora v_2 " è vero se e solo se tutti i Tipi di v_1 sono collegati con la relazione di prototipo o di istanza a una Scena prototipo di un Tipo di v_2 ;

- **Inferenza generale tra Tipo e verbo**

Dato un Tipo t_1 prototipo di una Scena s_1 e un verbo v_2 , l'enunciato "se t_1 allora v_2 " è vero se esiste un Tipo t_2 di v_2 che è prototipo di s_1 oppure che è prototipo di s_2 e s_1 è istanza di s_2 ;

- **Applicabilità di un verbo ad una Scena**

Dati una Scena s_1 e un verbo v_1 , v_1 è applicabile a s_1 se e solo se esiste un Tipo t_1 di v_1 tale che s_1 è un prototipo di t_1 oppure s_2 è un prototipo di t_1 e s_1 è istanza di s_2 ;

- **Equivalenza locale di due verbi**

Dati due verbi v_1 e v_2 , due tipi t_1 e t_2 appartenenti rispettivamente a v_1 e v_2 e una Scena s_1 prototipo di t_1 , v_1 è localmente equivalente a v_2 se e solo se esiste un Tipo t_2 di v_2 tale che s_1 è prototipo di t_2 oppure s_2 è prototipo di t_2 e s_1 è istanza di s_2 ;

- **Traducibilità di due verbi in lingue diverse**

Dati un verbo v_1 in una lingua L_1 e un verbo v_2 in una lingua L_2 diversa da L_1 , v_2 è un traduce di v_1 se e solo se v_1 e v_2 sono localmente equivalenti.

LIMITI DEL MODELLO GERARCHICO

In questo capitolo si mostra, attraverso specifici casi di studio, come il modello dati gerarchico sia inadatto a rappresentare correttamente i concetti azionali e – in generale – a rispondere ai requisiti richiesti per una risorsa linguistica come IMAGACT. Sono emersi infatti problemi relativi sia al modello inferenziale di questa struttura logica, sia alla sua adeguatezza per la rappresentazione coerente dei concetti. Inoltre dal lavoro sul mapping della lingua cinese nell'ontologia, è stata estratta una casistica di problematiche di traduzione, da cui sono emersi dei requisiti che il modello dati deve poter soddisfare per consentire l'estensione a nuove lingue.

5.1 ERRORI DI RAPPRESENTAZIONE NELLA STRUTTURA DATI

I casi di studio presentati di seguito derivano dall'analisi dei dati attualmente presenti in IMAGACT e ne evidenziano alcune criticità, che fanno capo a due categorie di problemi:

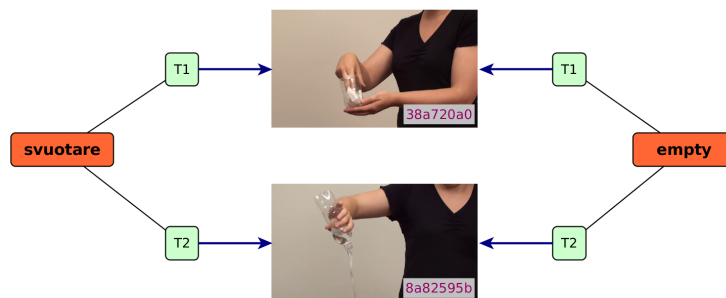
1. La creazione dei tipi non è indipendente dalla gerarchia delle scene: questo è particolarmente grave poiché contrasta con il modello teorico per cui la suddivisione in tipi azionali deve essere guidata dalla competenza semantica dell'annotatore madrelingua; invece talvolta l'annotatore è costretto ad operare scelte forzate, a causa dei limiti espressivi della struttura dati.
2. La rigidità logica rende impossibile la modifica dei dati: molto spesso, soprattutto nei casi in cui le scene sono incapsulate in alberi molto profondi, le modifiche ai dati sono difficili da gestire, poiché si ripercuotono su tutti i tipi collegati.

5.1.1 Caso 1. Costruzione gerarchica ad hoc

In alcuni casi sono presenti delle incoerenze nella rappresentazione della variazione verticale, che è stata alterata per permettere il collegamento di tipi diversi alla stessa Scena. Questa è una delle cause della dipendenza troppo forte della strutturazione in tipi dalla gerarchia di scene.

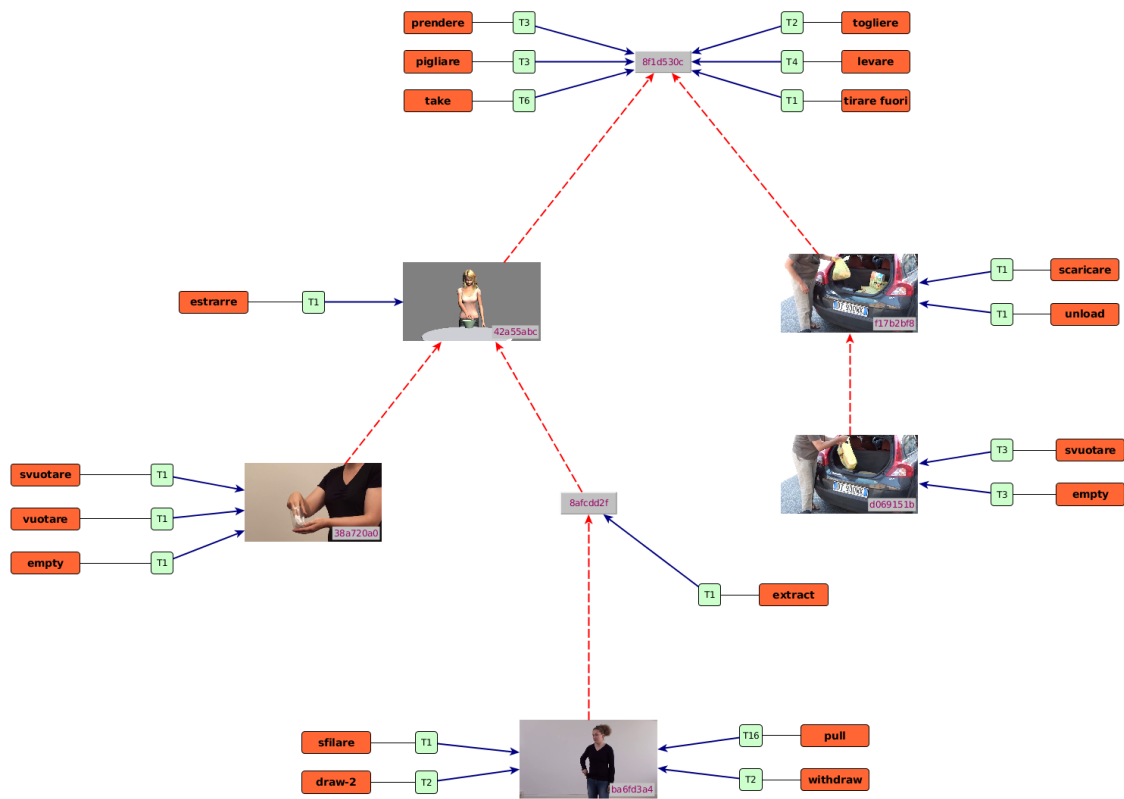
ESEMPIO 1: SVUOTARE E SCARICARE Nell'annotazione originale il verbo italiano *svuotare* e il verbo inglese *empty* erano stati suddivisi in due tipi azionali, differenziando gli eventi in cui il contenuto rimosso sia solido (*svuotare il barattolo delle penne*) oppure liquido (*svuotare la bottiglia dell'acqua*): questo tratto infatti è stato giudicato discriminativo di due concetti azionali distinti.

Figura 34: I due tipi di *svutare* e *empty*



Durante la fase di mapping è però emerso che il Tipo 1 (solido) poteva essere collegato sia ad un Tipo di *estrarre* (*estrarre la pallina dal bicchiere*), sia ad un Tipo di *scaricare* (*scaricare i pacchi dall'auto*); in entrambi questi eventi *svuotare-T1* può essere applicato nel caso particolare in cui il contenuto venga rimosso completamente.

In questo caso l'annotazione di occorrenze da corpus ha portato ad un'evidenza empirica difficilmente prevedibile, quella cioè che due verbi semanticamente distanti come *scaricare* e *svuotare* siano localmente equivalenti; in particolare essi sono equivalenti nel campo di variazione rappresentato dalla Scena "do69151b". Il Tipo *svuotare-T1* è pertanto stato suddiviso in due diversi tipi sulla base di un differenziale linguistico, che è il verbo *scaricare*. La presenza di un differenziale linguistico costituisce spesso un

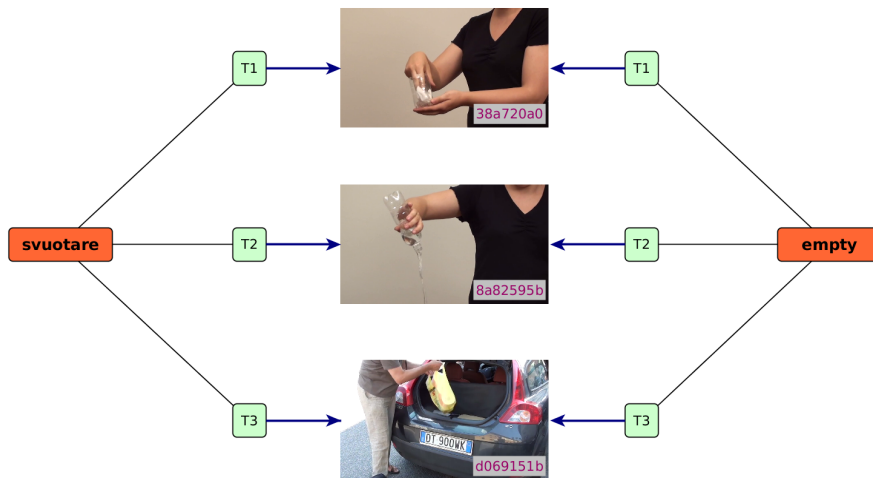
Figura 35: *svuotare* e *scaricare* nel modello gerarchico

fattore rilevante che giustifica la creazione di Tipo. Questo però non è valido in assoluto, e il caso corrente è un esempio in cui la presenza di un differenziale linguistico non è discriminativa della variazione del verbo *svuotare*, nel quale azioni come “Marco svuota la macchina” e “Marco svuota la credenza” vengono riconosciute tipologicamente analoghe.

Come mostrato in figura 35, i due tipi di *scaricare* ed *estrarre* non fanno parte della stessa tassonomia, cioè sono inseriti in due rami indipendenti della gerarchia e non esiste tra loro nessuna relazione di istanza. Pertanto, per poter collegare *svuotare-T1* come istanza ai due tipi di *estrarre* e *scaricare*, è stato necessario creare un altro Tipo di *svuotare*, discriminando il fatto che il contenitore sia un veicolo oppure no.

In questo modo è stato possibile fare il mapping tra *svuotare* e altri verbi “vicini” (come *estrarre*, *scaricare*, *prendere*, *togliere*,...) ma si è stati costretti ad aggiungere un Tipo artificiale con scarsa rilevanza pragmatica. È infatti poco rilevante, dal punto di vista del verbo *svuotare*, che l’oggetto svuotato sia un veicolo, un mobile o un bicchiere; mentre invece la natura liquida o solida del contenuto rimosso determina una spostamento

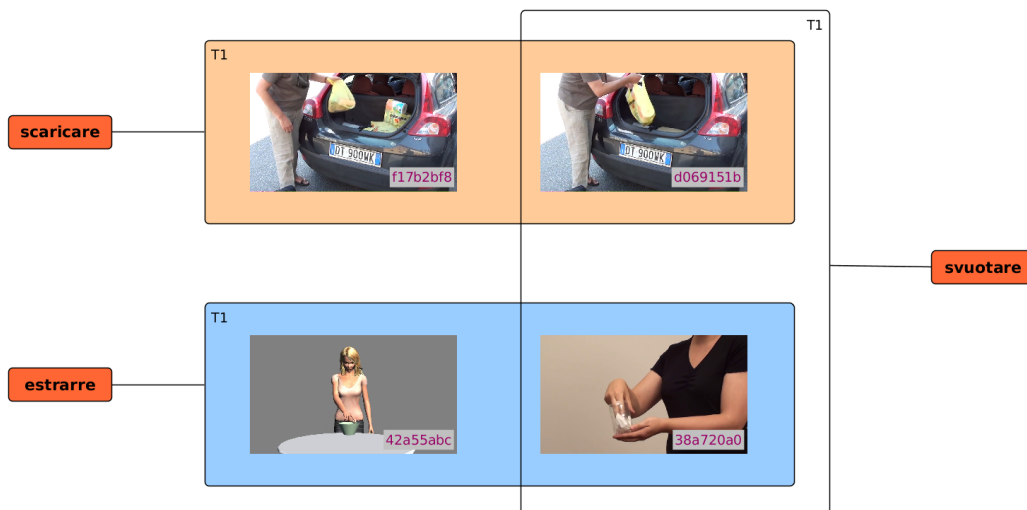
Figura 36: Aggiunta di un nuovo Tipo di *svuotare*



pragmatico che giustifica la creazione di due tipi azionali.

In questo esempio emerge chiaramente un limite espressivo del modello gerarchico, ossia l'incapacità di rappresentare l'intersezione insiemistica tra tipi che in casi come questo è necessaria, poiché permette di fare il mapping tra i verbi senza doverne alterare la suddivisione tipologica. Questo esempio è particolarmente interessante poiché mostra come la sovrapposizione si determini non soltanto tra i sensi delle parole, ma anche in partizioni di senso di livello molto più basso, come i Tipi.

Figura 37: L'intersezione tra i Tipi di *svuotare* e *scaricare*



In questo esempio il Tipo di *svuotare* potrebbe contenere due scene, una che fa parte del Tipo di *estrarre* e una che è collegata al Tipo di *scaricare*. Verrebbe così eliminato il vincolo gerarchico tra *estrarre-T₁*/*scaricare-T₁* e *svuotare-T₁* esprimendo correttamente la relazione che effettivamente intercorre tra questi tipi: essi non sono tra di loro più generali o più specifici (ovvero in relazione gerarchica), ma semplicemente condividono delle classi di eventi.

5.1.2 Caso 2. Mapping mancato tra tipi

Lo stesso problema nella difficoltà di effettuare un collegamento tra tipi diversi alla stessa Scena è stato in alcuni casi affrontato in modo opposto rispetto al caso 1, scegliendo cioè di non effettuare il mapping per non modificare la suddivisione in tipi azionali.

Figura 38: Il primo Tipo di *prendere*



ESEMPIO 2: PRENDERE E TOGLIERE La Scena collegata al primo Tipo di *prendere* può essere descritta da due punti di vista, uno che focalizza sull'atto di presa dell'oggetto e l'altro sulla rimozione volontaria dell'oggetto dalla sua locazione. Questa dualità è esprimibile in italiano con due verbi, *prendere* ("Fabio prende la tazza dal ripiano") e *togliere* ("Fabio toglie la tazza dal ripiano"), entrambi perfettamente adeguati per descrivere la Scena. Inoltre in azioni di questo Tipo abbiamo una relazione di implicazione tra *prendere* e *togliere*: se "Fabio prende la tazza dal ripiano" allora "Fabio toglie la tazza dal ripiano"; mentre non è vero il viceversa (Fabio può togliere la tazza in molti modi senza prenderla, ad esempio può spingerla fuori dal ripiano). Questi fatti si traducono in IMAGACT con un collegamento della Scena sia ad un Tipo di *togliere* (con la relazione di **prototipo**), sia ad un Tipo di *prendere* (con la relazione di **istanza**); si evidenzia così

una maggiore generalità del Tipo di *togliere* rispetto al Tipo di *prendere*, coerentemente con ciò che emerge dall'esempio. Avremo quindi un Tipo di *prendere* - che potremmo sommariamente descrivere come "prendere un oggetto solido da una locazione" - che è un caso particolare del Tipo di "togliere un oggetto solido da una locazione", più generale in quanto categorizza anche eventi in cui l'oggetto viene tolto senza essere preso.

Come vediamo dalla figura 38, però, la Scena non è collegata a *togliere* nell'ontologia (il verbo *togliere* non è elencato tra quelli applicabili alla Scena); questa mancanza è dovuta ad un problema che sta dietro all'inserimento degli oggetti in questa struttura dati gerarchica: per collegare questa Scena a entrambi i verbi, infatti, dovremmo necessariamente assumere che la "rimozione dell'oggetto" sia un tratto così importante dal punto di vista del verbo *prendere* da giustificare la creazione di un Tipo. In altre parole per poter avere un Tipo di *prendere* che è più specifico di un Tipo di *togliere* c'è bisogno che questo Tipo di *prendere* categorizzi solo eventi che sono anche *togliere* e che quindi la "rimozione dell'oggetto" sia un tratto distintivo e necessario del Tipo. Questo si scontra con i criteri di sviluppo dell'ontologia, in cui i tipi vengono determinati non sulla base di una discriminazione basata su tratti semantici predefiniti, ma dall'analisi dell'annotazione di esempi estratti da corpus. In questo Tipo di *prendere*, la rimozione dell'oggetto non è stato considerato un tratto importante in fase di annotazione, infatti al suo interno troviamo sia frasi in cui *prendere* è anche *togliere* (come "Fabio prende la tazza dal ripiano"), sia frasi in cui *prendere* non è *togliere* come "Marco prende la penna". La presenza della locazione dell'oggetto, che in questo caso rappresenta la *sorgente* dell'azione, è stato considerato soltanto come un fattore di variazione argomentale, che non determina una variazione tipologica. Osservando l'amplessima variazione verticale di *prendere* appare chiaro che inserire "Fabio prende la tazza dal ripiano" e "Marco prende la penna" in due diversi tipi azionali sarebbe stato inadeguato. Per non introdurre questa criticità è stato scelto di non fare il mapping tra *prendere* e *togliere*, lasciando una evidente incompletezza nell'ontologia. Come si vede dalla seguente figura, un Tipo di *togliere* potrebbe accogliere tranquillamente come sua istanza la Scena di figura 38; la seconda Scena, infatti, è quasi identica a quella di *prendere*.

Questo caso è inverso rispetto a quello di *scaricare-svuotare*, nel quale il dato empirico del corpus ha fatto emergere un campo di variazione comune tra due verbi apparentemente

Figura 39: Il terzo Tipo di *togliere*

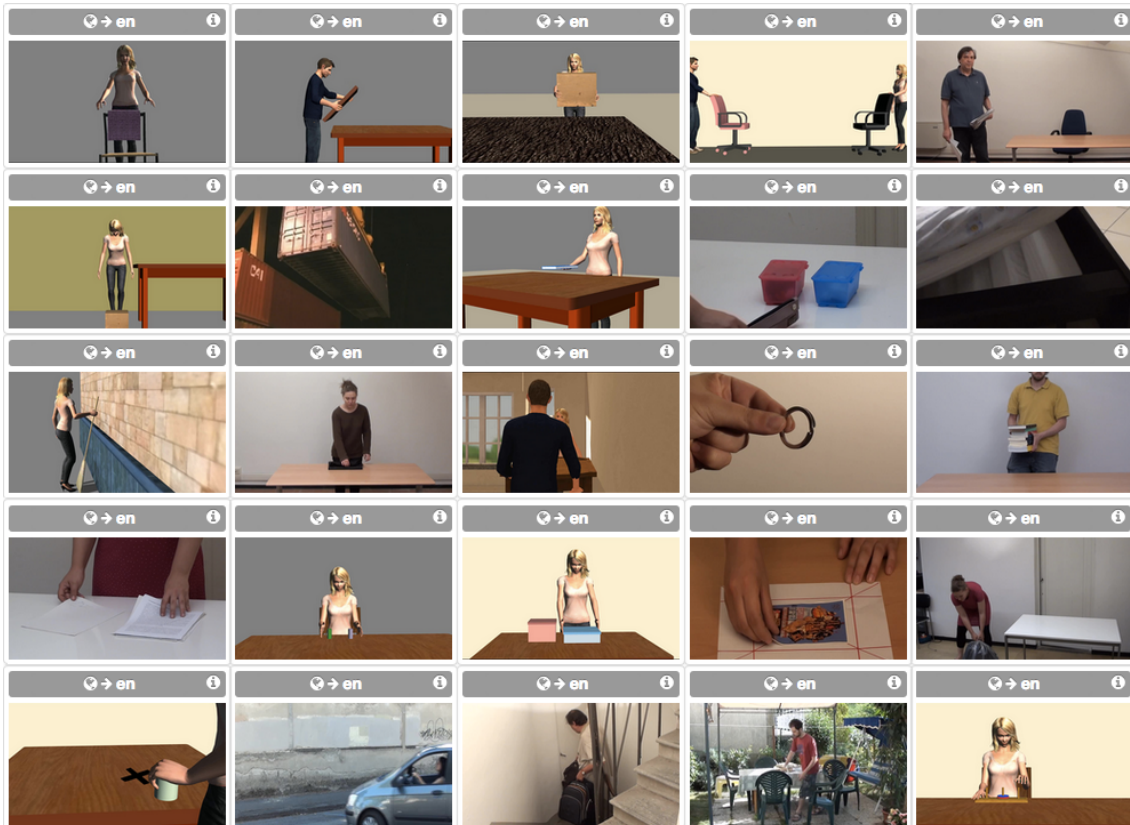
disgiunti: lì il differenziale linguistico, cioè l'applicabilità di *scaricare* ad alcune istanze di *svuotare*, ha portato ad indurre una suddivisione tipologica più fine. Nell'esempio corrente invece abbiamo una interdipendenza chiara tra i due tipi azionali di *prendere* e *togliere*, in cui l'inferenza "se prendo il bicchiere dal tavolo allora tolgo il bicchiere dal tavolo" sembra logicamente valida sull'intero Tipo di *prendere*. In questo caso però le evidenze del corpus smentiscono questa osservazione, mostrando delle occorrenze di *prendere* che fanno parte dello stesso Tipo, ma che non sono *togliere*, come ad esempio "Marco prende la penna". Per poter collegare *prendere* e *togliere* avremmo quindi bisogno di tre scene differenziali, (a) una in cui si prende e si toglie ("Marco prende la tazza dal ripiano"), (b) una in cui si prende ma non si toglie ("Marco prende la penna") e (c) una in cui si toglie e non si prende ("Marco toglie il bicchiere dal tavolo spingendolo"). In tal modo potremmo creare una relazione di *istanza* tra (a) e (c), un Tipo di *prendere* prototipo di (a) e un altro Tipo di *prendere* prototipo di (b). Questo però porterebbe ad avere due tipi di *prendere* troppo simili tra di loro e difficilmente differenziabili da un punto di vista pragmatico. Anche in questo caso viene evidenziato il vincolo troppo rigido che la gerarchia di scene impone sulla determinazione dei tipi azionali: se nel caso 1 il limite era stato superato introducendo una forzatura (con l'aggiunta di un Tipo artificiale), qui è stato scelto di preservare la categorizzazione tipologica lasciando però il mapping incompleto.

5.1.3 Caso 3. Difficoltà di modifica per tipi molto generali

Negli esempi precedenti abbiamo visto come la dipendenza tra tipi azionali sia un ostacolo al mapping dei tipi con le scene (casi 1 e 2); in questo esempio è mostrato come le relazioni rigide della struttura dati siano un ostacolo anche per la modifica di tipi

esistenti, in particolare per quelli collegati a molte scene.

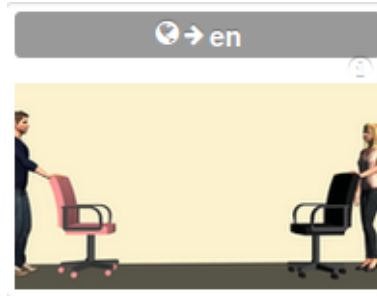
Figura 40: Il primo Tipo di *mettere*



ESEMPIO 3: METTERE Il Tipo *mettere-T1* è molto generale e racchiude una famiglia di 25 scene, alcune delle quali però designano azioni molto diverse tra loro e non classificabili in un unico concetto pragmatico.

Ad esempio la Scena 41, in cui due sedie vengono avvicinate, è difficilmente assimilabile alle altre scene di figura 40, dove gli oggetti vengono tutti per lo più appoggiati.

Fenomeni di questo Tipo sono comuni nei tipi generali, in cui la variazione è talmente ampia da rendere difficile la categorizzazione; inoltre la natura semantica dei tipi non è sempre chiarificata dall'insieme delle scene che ne fanno parte ed i tipi molto generali soffrono spesso di problemi di difficoltà interpretativa, come abbiamo detto nel paragrafo 3.3. In generale comunque la struttura dei dati dovrebbe essere sufficientemente flessibile da consentire di correggere le aberrazioni; in realtà però questo è estremamente complicato e faticoso nei tipi generali, in cui una modifica si ripercuote

Figura 41: Prototipo di un Tipo di *avvicinare*

su tutti gli elementi della gerarchia, ciascuno dei quali deve essere poi revisionato e, eventualmente, modificato per poter mantenere la coerenza logica. In questo esempio *mettere-T1* è incapsulato in una gerarchia di 60 tipi azionali, almeno 20 dei quali dovrebbero essere modificati se soltanto si decidesse di spostare la Scena di Fig. 41 in un altro Tipo.

I casi di studio presentati mostrano in modo evidente che una struttura dati basata sul modello gerarchico, sebbene si presti bene alla formulazione di inferenze logiche, è troppo rigida per modellare il linguaggio naturale e porta a rendere complicate (o addirittura impossibili) operazioni di base come inserimento, modifica e mapping. Poter fare queste operazioni in modo semplice è necessario per la gestione e la revisione dell'ontologia e pertanto costituisce un requisito funzionale della struttura dati che, sulla base delle problematiche analizzate, dovrà essere sufficientemente espressiva da permettere di:

- esplicitare i rapporti tra tipi attraverso intersezione insiemistica;
- inserire una Scena che sia il riferimento di due tipi semanticamente diversi;
- togliere o aggiungere una Scena da un Tipo senza modificarne le proprietà.

L'estensione di IMAGACT alla lingua cinese ci ha permesso di evidenziare dei punti critici relativi alla traduzione dei verbi generali tra due lingue lontane; in particolare sono emersi problemi di traducibilità, che in alcuni casi mettono in discussione il carattere

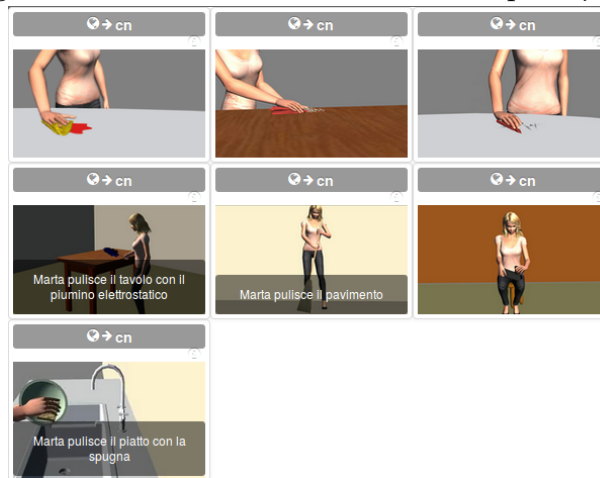
universale della rappresentazione del concetto attraverso un prototipo. Il diverso modo di categorizzare gli eventi in due lingue come l'italiano e il cinese, porta infatti ad avere diverse interpretazioni delle scene. Sebbene questo fenomeno sia circoscritto ad un numero esiguo di concetti azionali, che in larga parte sono condivisi tra le due lingue, esso ha portato ad un ripensamento del modello teorico, in particolare sui concetti di Tipo e Scena e sui rapporti tra queste due entità. Questa analisi è riassunta in 6 casi di studio, ciascuno dei quali fa riferimento ad uno specifico problema di traduzione.

5.2.1 Caso 1. Mancanza di un verbo generale

All'interno di IMAGACT, e specificamente nella lingua cinese, abbiamo documentato una serie di casi in cui manca un verbo generale in grado di applicarsi a tutte o a una buona parte delle azioni categorizzabili con un verbo generale italiano o inglese. Notiamo che questo è abbastanza frequente nella lingua cinese, in cui i verbi generali sono molti meno rispetto a quelli presenti nelle lingue europee. Questo fatto non è paragonabile alla differenza di estensione che hanno due verbi generali di due diverse lingue, come ad esempio *mettere* in italiano e *put* in inglese: essi infatti, benché non condividano tutta la variazione verticale, hanno molti tipi che, con un certo grado di approssimazione, sono equivalenti, cioè coprono le stesse classi di azione ed hanno un livello simile di generalità. Con la lingua cinese invece capita abbastanza frequentemente che un Tipo che in italiano o in inglese corrisponde ad un concetto coeso identificato da tratti semantici prominenti, non sia categorizzato e quindi non esista un verbo cinese in grado di applicarsi a tutta la sua estensione (o a una buona parte di essa). Questo fenomeno linguistico, noto in letteratura come *lexical gap*, è stato riscontrato varie volte anche nel mapping italiano-inglese, ma sempre per piccole variazioni semantiche; in questi casi della lingua cinese invece il *gap* è particolarmente significativo e riguarda concetti generali. Sebbene questo fenomeno abbia un impatto ridotto, dal momento che un'alta percentuale di concetti sono produttivi interlinguisticamente anche tra l'italiano e il cinese (vedi Par. 8.1), questi casi sono interessanti da analizzare poiché riguardano appunto verbi italiani ad alta frequenza e generali. Di seguito riportiamo due esempi di questo fenomeno, *pulire* e *suonare*.

ESEMPIO 4: PULIRE Il Tipo 1 di *pulire* si riferisce ad azioni pragmaticamente molto diverse (come spazzare per terra, spolverare il tavolo, dare il cencio, lavare un piatto,...) che però hanno la stessa finalità: togliere lo sporco. Questo tratto è molto forte per il verbo *pulire* e consente di raggruppare tutte queste azioni all'interno di un unico concetto semantico; un parlante competente difficilmente identificherà questa variazione come una polisemia del verbo *pulire*, ma piuttosto come tanti modi di fare la stessa azione.

Figura 42: Variazione orizzontale di un Tipo di *pulire*



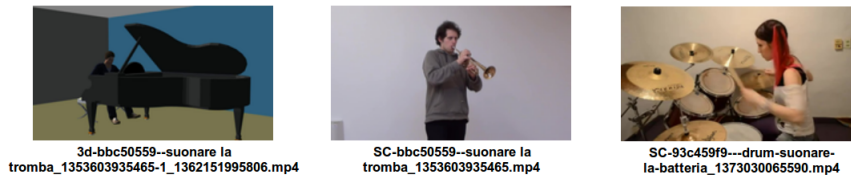
Questo concetto generale di *pulire* non può essere predicato in cinese, in cui esistono molti verbi specifici che identificano le singole azioni, ma non c'è un verbo che le categorizza tutte sulla base della loro finalità. Ad esempio per descrivere l'azione di "pulire il tavolo con il piumino elettrostatico" si usa il verbo 掸灰尘 (dǎn huī chén), per "pulire il pavimento" si usa 扫 (sǎo) e per "pulire il piatto" si usa 擦 (cā); non è però possibile riferirsi a tutte queste azioni con un solo verbo, come invece si può fare in italiano con una frase come "Marco pulisce".

Nella struttura dati gerarchica, per poter fare il mapping, i tipi dei vari verbi cinesi 掸灰尘 (dǎn huī chén), 扫 (sǎo), 擦 (cā),... devono essere collegati al Tipo di *pulire* con una relazione di istanza. Per fare ciò bisogna imporre il vincolo che i verbi cinesi coinvolti abbiano un Tipo definito sul concetto di "pulizia" che sia quindi un'istanza di *pulire*; non è detto però che questa suddivisione sia sensata per il verbo cinese in questione. Il verbo 擦 (cā), ad esempio, predica di un'azione eseguita per sfregamento; per poterlo mappare con *pulire* è necessario che 擦 (cā) abbia almeno due tipi, uno in cui vengono classificate

le azioni in cui lo sfregamento comporta una pulizia e una in cui c'è sfregamento senza pulizia. Come succedeva anche nel caso d'uso di prendere/togliere, appare evidente il forte limite di questa strutturazione dell'informazione in cui la creazione di tipi azionali non può essere fatta liberamente sulla base della competenza semantica, ma deve sottostare a vincoli imposti dalla gerarchia.

ESEMPIO 5: SUONARE Non c'è un verbo cinese adatto a descrivere le azioni riferibili con il Tipo *suonare-T1* (che corrisponde a “suonare uno strumento musicale”), ma una molteplicità di verbi specifici per lo strumento suonato.

Figura 43: Applicabilità dei verbi al primo Tipo di *suonare*



	★ 吹 chuī	★ 弹 tán	★ 拉 lā	★ 打 dǎ	★ 击 jī	★ 演奏 yǎn zòu
Matteo suona la tromba	Y: <input checked="" type="checkbox"/> N: <input type="checkbox"/>	Y: <input type="checkbox"/> N: <input type="checkbox"/> P: <input type="checkbox"/>	Y: <input type="checkbox"/> N: <input checked="" type="checkbox"/> P: <input type="checkbox"/>	Y: <input type="checkbox"/> N: <input checked="" type="checkbox"/> P: <input type="checkbox"/>	Y: <input type="checkbox"/> N: <input checked="" type="checkbox"/> P: <input type="checkbox"/>	Y: <input type="checkbox"/> N: <input checked="" type="checkbox"/> P: <input type="checkbox"/>
<input type="checkbox"/> Michele suona il pianoforte (1)	Y: <input type="checkbox"/> N: <input checked="" type="checkbox"/>	Y: <input checked="" type="checkbox"/> N: <input type="checkbox"/> P: <input checked="" type="checkbox"/>	Y: <input type="checkbox"/> N: <input checked="" type="checkbox"/> P: <input type="checkbox"/>	Y: <input type="checkbox"/> N: <input checked="" type="checkbox"/> P: <input type="checkbox"/>	Y: <input type="checkbox"/> N: <input checked="" type="checkbox"/> P: <input type="checkbox"/>	Y: <input type="checkbox"/> N: <input checked="" type="checkbox"/> P: <input type="checkbox"/>
<input type="checkbox"/> Matteo suona la tromba (1)	Y: <input checked="" type="checkbox"/> N: <input type="checkbox"/>	Y: <input type="checkbox"/> N: <input checked="" type="checkbox"/> P: <input type="checkbox"/>	Y: <input type="checkbox"/> N: <input checked="" type="checkbox"/> P: <input type="checkbox"/>	Y: <input type="checkbox"/> N: <input checked="" type="checkbox"/> P: <input type="checkbox"/>	Y: <input type="checkbox"/> N: <input checked="" type="checkbox"/> P: <input type="checkbox"/>	Y: <input type="checkbox"/> N: <input checked="" type="checkbox"/> P: <input type="checkbox"/>
<input type="checkbox"/> Michele suona il sassofono (1)	Y: <input checked="" type="checkbox"/> N: <input type="checkbox"/>	Y: <input type="checkbox"/> N: <input checked="" type="checkbox"/> P: <input type="checkbox"/>	Y: <input type="checkbox"/> N: <input checked="" type="checkbox"/> P: <input type="checkbox"/>	Y: <input type="checkbox"/> N: <input checked="" type="checkbox"/> P: <input type="checkbox"/>	Y: <input type="checkbox"/> N: <input checked="" type="checkbox"/> P: <input type="checkbox"/>	Y: <input type="checkbox"/> N: <input checked="" type="checkbox"/> P: <input type="checkbox"/>
<input type="checkbox"/> Luigi suona il flauto (1)	Y: <input checked="" type="checkbox"/> N: <input type="checkbox"/>	Y: <input type="checkbox"/> N: <input checked="" type="checkbox"/> P: <input type="checkbox"/>	Y: <input type="checkbox"/> N: <input checked="" type="checkbox"/> P: <input type="checkbox"/>	Y: <input type="checkbox"/> N: <input checked="" type="checkbox"/> P: <input type="checkbox"/>	Y: <input type="checkbox"/> N: <input checked="" type="checkbox"/> P: <input type="checkbox"/>	Y: <input type="checkbox"/> N: <input checked="" type="checkbox"/> P: <input type="checkbox"/>
<input type="checkbox"/> Nello suona la chitarra (1)	Y: <input type="checkbox"/> N: <input checked="" type="checkbox"/>	Y: <input checked="" type="checkbox"/> N: <input type="checkbox"/> P: <input checked="" type="checkbox"/>	Y: <input type="checkbox"/> N: <input checked="" type="checkbox"/> P: <input type="checkbox"/>	Y: <input type="checkbox"/> N: <input checked="" type="checkbox"/> P: <input type="checkbox"/>	Y: <input type="checkbox"/> N: <input checked="" type="checkbox"/> P: <input type="checkbox"/>	Y: <input type="checkbox"/> N: <input checked="" type="checkbox"/> P: <input type="checkbox"/>

Come si vede in Fig. 43 non c'è un verbo cinese che copra l'intera variazione orizzontale del primo Tipo di suonare, ma si usa 弹 (tán) per il pianoforte e la chitarra, 击 (jī) e 打 (dǎ) per le percussioni, 吹 (chuī) per gli strumenti a fiato, 拉 (lā) per gli archi; c'è poi il verbo 演奏 (yǎn zòu) che è applicabile a tutti gli strumenti, ma significa specificamente “suonare in pubblico, fare una performance”. Come per pulire quindi i tratti che in italiano sono rilevanti (in questo caso “fare musica attraverso uno strumento”) e creano un concetto generale, in cinese non lo sono; per cui le azioni vengono riferite da verbi più specifici che focalizzano sull'azione fisica compiuta (premere, battere, soffiare,...) anziché sull'azione più astratta di “fare musica”. In tutti questi casi in cui in una lingua non

c'è un corrispettivo verbo generale, abbiamo problemi nella traduzione di frasi in cui il verbo generale è usato in modo astratto, come “Marco suona” o “Marco pulisce”.

5.2.2 Caso 2. Riferimento all'azione da diversi punti di vista

In alcuni casi la stessa azione in cinese e in italiano è riferita da due diversi punti di vista; quando ciò accade abbiamo due verbi semanticamente diversi, che però si riferiscono alla stessa azione.

ESEMPIO 6: ABBASSARE L'azione rappresentata in Fig. 44 (*abbassare il cofano dell'auto*) è predicabile in italiano con il verbo *abbassare* e in cinese con il verbo 盖 (gài), che significa “coprire”. I verbi *abbassare* e 盖 (gài) non hanno però una semantica vicina, tanto che è possibile utilizzare 盖 (gài) per azioni in cui si copre un oggetto senza abbassare il coperchio (ad esempio una scatola con il coperchio a scorrimento). Viceversa è possibile applicare *abbassare* in azioni pragmaticamente molto simili a questa, ma in cui non ci sia niente da coprire (banalmente possiamo pensare ad abbassare il cofano fino a metà), e quindi non predicabili con 盖 (gài).

Figura 44: Azione “abbassare il cofano dell'auto”



In casi come questo non abbiamo alcuna relazione semantica tra i due verbi, ma un riferimento comune alla Scena. È importante notare che questo esempio non è relativo a usi periferici di lemmi verbali, ma 盖 (gài) è il verbo che un parlante di lingua cinese utilizza normalmente per riferirsi a questa azione; anzi, in questo caso è il solo verbo possibile. Le conseguenze di ciò sono forti dal punto di vista della traduzione: stiamo infatti asserendo che in questo contesto pragmatico l'unico traduttore cinese possibile per il verbo italiano *abbassare* è un verbo il cui significato non è semanticamente correlato

ad *abbassare*. In questo esempio appare chiara l'importanza del contributo che IMAGACT può fornire alla traduzione: attraverso l'utilizzo del riferimento azionale infatti, vengono identificati possibili traducanti, altrimenti difficili da prevedere.

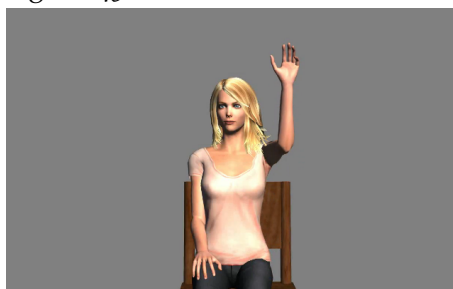
5.2.3 Caso 3. Azioni senza scopo

Nella lingua cinese lo scopo per cui si compie un'azione è spesso un tratto necessario per poterla predicare: questo perché per molti verbi cinesi il focus dell'azione è sul risultato.

ESEMPIO 7: ALZARE LA MANO In questo esempio si fa riferimento al Tipo 8 del verbo italiano *alzare*, identificato dalla Scena "Marta alza la mano" (Fig. 45). Questa azione non può essere predicata in cinese senza conoscerne lo scopo. Dovremmo ad esempio utilizzare:

- il verbo 举 (jǔ) , nei casi in cui si alzano le mani per arrendersi al nemico;
- il verbo 伸 (shēn) , quando si alza la mano per prendere un oggetto che si trova in alto;
- il verbo 举手 (jǔ shǒu) , se si alza la mano per votare o per prendere la parola.

Figura 45: Azione "alzare la mano"



Non soltanto la stessa azione con diversi scopi richiede l'uso di verbi diversi, ma non possiamo utilizzare un verbo più generale per predicare di quell'azione senza conoscerne lo scopo. Ne segue che:

1. la Scena rappresentata in Fig. 45 non è descrivibile sinteticamente in cinese;

2. il Tipo 8 di alzare non è traducibile con un solo verbo cinese; ci sono però almeno tre verbi che possono essere candidati alla traduzione.

5.2.4 Caso 4. Azioni senza verbo

Non sempre un'azione fisica è predicabile con un verbo: in alcuni casi per riferirvisi è necessario utilizzare una perifrasi in cui il verbo non c'è oppure è presente, ma soltanto come supporto e il concetto azionale è espresso da strutture sintattiche senza verbo.

ESEMPIO 8: AVVICINARE La classe di azioni rappresentata dal prototipo di Fig. 46 (*Marta avvicina i due cilindri*) non viene descritta in cinese attraverso un verbo, come invece avviene in italiano con *avvicinare*, ma attraverso una perifrasi: 让两个物体靠近一些 (ràng liǎng gè wù tǐ kào jìn yì xiē). In questa frase l'unico verbo presente è 让 (ràng), che però ha solo un ruolo di supporto e non esprime concetti azionali. Il significato di 让 (ràng) è simile a quello dell'inglese *to let* (fare sì che..., fare in modo che...); una traduzione letterale di questa frase potrebbe essere "fare in modo che gli oggetti siano vicini".

Figura 46: Azione "avvicinare due cilindri"



Potremmo dire che il concetto di *avvicinare* in cinese non è codificato tramite un verbo; pertanto qui l'ontologia IMAGACT, in cui l'unità linguistica è il verbo, non è informativa per la traduzione. Infatti la Scena di Fig. 46 non sarà legata ad alcun verbo cinese. Per sopperire a questa mancanza è stata data la possibilità all'annotatore di inserire perifrasi al posto di verbi, in modo da avere comunque un output di traduzione.

5.2.5 Caso 5. Movimento fuori dal verbo

In cinese il movimento spesso non è codificato con il verbo, ma con preposizioni o altre parti del discorso; viceversa in italiano sono frequenti verbi che esprimono un concetto di movimento: *spostare*, *portare*, *alzare*, ecc. In molti casi questa discrepanza porta ad avere non solo un'altra struttura di predicazione, ma anche una diversa concettualizzazione dell'azione, in cui il verbo fa riferimento a differenti concetti azionali.

ESEMPIO 9: PORTARE Un'azione come quella presentata in Fig. 47, "Marta porta la valigia", in cui il verbo italiano *portare* codifica il movimento, viene espressa in cinese dalla frase 玛尔塔拎着行李箱 (mǎ ěr tǎ līn zhe xíng lǐ xiāng), in cui il verbo cinese 拎 (līn) non codifica il movimento, bensì il modo di tenere l'oggetto (in questo caso "in mano").

Figura 47: Azione "portare una valigia"



Notiamo quindi che il modo di tenere l'oggetto (in braccio, sulle spalle, in mano, ...) è un tratto essenziale per poter descrivere l'azione in cinese, poiché da esso dipende il verbo che deve essere utilizzato, come negli esempi seguenti:

- John porta la tv (in braccio) in soggiorno → 约翰把电视机抱到客厅 (yuē hàn bǎ diàn shì jī bào dào kè tīng) .
- John porta il sacco di riso (sulle spalle) in cucina → 约翰把一袋大米扛到厨房 (yuē hàn bǎ yí dài dà mǐ káng dào chú fáng) .
- John porta la busta della spesa a casa → 约翰把买的东西拎到家里 (yuē hàn bǎ mǎi de dōng xi līn dào jiā lǐ) .

- John porta la penna sul tavolo → 约翰把钢笔拿到桌上 (yuē hàn bǎ gāng bǐ ná dào zhuō shàng)

Se da un lato non c'è alcuna difficoltà a descrivere in cinese una Scena come quella in Fig. 47, in cui questo tratto è chiaramente esplicitato, dall'altro sussiste un problema di traduzione, poiché il verbo *portare* non predica del modo di tenere l'oggetto. Anzi, poiché questo è un tratto irrilevante dal punto di vista di *portare*, possiamo aspettarci che in molti casi non venga esplicitato, rendendo quindi impossibile la traduzione dell'azione in cinese.

5.2.6 Caso 6. Mancanza del concetto azionale

I verbi sono la parte del discorso che normalmente è preposta a categorizzare gli eventi; questi però non sono sempre riferibili da un unico verbo: un evento complesso in cui accadono diverse azioni in sequenza può essere scomposto in una serie di eventi semplici per poter essere descritto. La complessità però non è indipendente dalla lingua, o almeno non lo è completamente: alcuni eventi che in una lingua possono essere descritti con un solo verbo, necessitano invece di più verbi in un'altra lingua, poiché devono essere suddivisi in eventi più semplici.

ESEMPIO 10: MUOVERE In questo esempio l'azione rappresentata può essere descritta in italiano dal verbo generale *muovere* con una frase semplice: "Marta muove le gambe". La lingua cinese invece, non ha un verbo in grado di riferirsi unitariamente questa azione, che infatti non viene neanche riconosciuta come un'azione, ma come una sequenza di azioni. Un parlante cinese potrebbe predicare di questo evento con 把脚抬起来, 上下晃动 (bǎ jiǎo tái qǐ lái, shàng xià huàng dòng), che corrisponde all'incirca a "alzare i piedi e farli oscillare su e giù"; in questa descrizione ci sono due verbi, 抬 (tái) e 晃动 (huàng dòng), rispettivamente traducibili come *alzare* e *oscillare*.

Casi come questo sono abbastanza rari, ma costituiscono un problema di traduzione difficile da risolvere: se nel Caso 4 infatti si poteva utilizzare una perifrasi per esprimere un concetto azionale, qui manca l'identificazione stessa dell'azione e l'unico modo per descrivere ciò che accade nella Scena è attraverso la composizione di più frasi. Da un

Figura 48: Azione “muovere le gambe”



punto di vista operativo appare quindi poco sensato legare una traduzione (un verbo o una perifrasi) cinese a questa Scena, che, sottoposta ad un madrelingua cinese, viene giudicata strana e poco chiara.

5.3 NECESSITÀ EMPIRICHE DERIVANTI DAI CASI DI STUDIO

Questi casi di studio, che mettono in evidenza alcune problematiche sulla possibilità di tradurre i verbi generali in una lingua lontana dall'italiano come il cinese, determinano anche delle qualità che il nuovo modello dati deve avere per rendere possibile l'estensione dell'ontologia ad altre lingue.

5.3.1 *Aggiungere nuovi tipi*

Il primo requisito, che la struttura gerarchica non ha, è quello di consentire l'inserimento di nuovi tipi creati da verbi di altre lingue. Per adesso infatti in IMAGACT solo l'italiano e l'inglese hanno la suddivisione in tipi azionali, mentre tutte le altre lingue hanno un mapping diretto tra il verbo ed il suo riferimento (la Scena). Collegare verbi a scene è un'operazione compatibile con l'attuale struttura dati, mentre la suddivisione in tipi è un punto problematico proprio a causa del modello gerarchico che sottosta all'ontologia. Quando si inserisce un nuovo Tipo nella struttura dati ci sono 5 possibili casi:

1. Il nuovo Tipo è “equivalente” ad un Tipo presente in gerarchia: in questo caso i due tipi rappresentano la stessa classe di azioni e, pertanto, categorizzano lo stesso set di scene. Il mapping si risolve con l'aggiunta di una relazione di prototipo tra i

due tipi;

2. Il nuovo Tipo è più specifico di un Tipo presente in gerarchia: in questo caso la serie di azioni coperte dal nuovo Tipo è più ristretta rispetto a quelle del Tipo nell'ontologia; il mapping può essere fatto collegando i due tipi con una relazione di istanza e aggiungendo al nuovo Tipo una Scena che sia prototipica e differenziale rispetto all'altro Tipo;
3. Il nuovo Tipo è più generale di un Tipo presente in gerarchia: in questo caso il nuovo Tipo categorizza una serie più ampia di azioni rispetto al Tipo dell'ontologia; per fare il mapping è necessario collegare i due tipi con una relazione di istanza e aggiungere al nuovo Tipo una Scena che sia prototipica e differenziale rispetto all'altro Tipo e a tutti i tipi figli;
4. Il nuovo Tipo è più generale di due o più tipi presenti in gerarchia: in questo caso il nuovo Tipo rappresenta una classe di azioni estesa che copre tutte le azioni degli altri tipi dell'ontologia; per fare il mapping è necessario collegare i tipi dell'ontologia al nuovo Tipo con una relazione di istanza; non è necessario aggiungere una Scena, poiché le scene degli altri tipi costituiscono già un differenziale;
5. Il nuovo Tipo non è né equivalente, né più specifico, né più generale di altri tipi presenti in gerarchia, ma fa riferimento ad un diverso concetto semantico e copre azioni che non sono collegate in gerarchia o che sono collegate ma a livelli diversi; in questo caso il mapping non è possibile senza una modifica ai tipi: è infatti necessario adeguare il nuovo Tipo al modello dati (dividendolo in più sottotipi che si possano inserire in gerarchia) oppure adeguare l'ontologia al nuovo Tipo (operando modifiche sui tipi già mappati).

Se nei primi 4 casi è possibile gestire facilmente l'inserimento di un nuovo verbo, il quinto caso rappresenta un problema che non può essere trattato con l'attuale struttura dati. Per farlo infatti è necessario che l'inserimento di un Tipo nell'ontologia sia indipendente dalla semantica del Tipo stesso.

Come abbiamo visto nel caso 1, attualmente non è così: la creazione dei tipi azionali infatti non può essere fatta soltanto sulla base della competenza dell'annotatore, ma deve

sottostare ai vincoli imposti dalla gerarchia nella quale il Tipo sarà mappato. In altre parole, si deve poter collegare un Tipo di 打 (dǎ) alla Scena “Marta suona la batteria” senza dover imporre che quel Tipo esiste in virtù del fatto che 打 (dǎ) è anche “suonare”. Staremmo in questo caso definendo un Tipo sulla base di un differenziale linguistico di un'altra lingua, anziché della propria. Questo è profondamente sbagliato, dal momento che la categorizzazione linguistica è diversa tra lingua e lingua, come peraltro emerge chiaramente dai casi appena descritti: suonare il violino o la batteria non sono due attività che necessitano diversa categorizzazione a livello linguistico e cognitivo per un italiano.

Dal caso 2 otteniamo un vincolo ancora più forte: deve essere possibile collegare alla stessa Scena due tipi tra cui non sussiste alcuna relazione semantica. Soltanto così infatti si può mappare insieme *abbassare* e 盖 (gài), che, pur non avendo un significato simile, si estendono entrambi ad una stessa classe di azioni. Deriviamo da qui due condizioni necessarie per consentire l’inserimento di nuovi tipi da altre lingue:

- i tipi devono poter essere definiti in modo indipendente dagli altri;
- l’equivalenza locale di due verbi non deve necessariamente implicare una relazione semantica tra essi.

5.3.2 Aggiungere nuove scene

Il secondo requisito, che deriviamo dal caso 3, è la possibilità di inserire nuove scene all’ontologia, quando quelle presenti non siano sufficientemente rappresentative. Devono essere ad esempio inserite scene più specifiche di “Marta alza la mano”, che esplicitino lo scopo dell’azione; queste in italiano arricchiranno la rappresentazione del verbo *alzare*, mentre in cinese saranno prototipi azionali dei verbi 举 (jǔ), 伸 (shēn) e 举手 (jǔ shǒu), che non possono invece estendersi all’azione più generale. In questo caso le nuove scene contengono un tratto semantico in più e sono quindi più specifiche dal punto di vista del verbo italiano *alzare*: in accordo con le definizioni del modello gerarchico di IMAGACT, queste scene sono istanze del prototipo azionale di un Tipo di alzare.

La necessità di inserire nuove scene la si trova anche nel caso 5; esse però non si realizzano con l'aggiunta di tratti semantici, ma con una modificazione degli stessi: se la Scena di "Marta alza la mano" infatti non è riferibile da un verbo cinese perché non esplicita lo scopo dell'azione, la Scena di "Marta porta la valigia" è invece perfettamente descrivibile perché contiene il modo con cui l'oggetto viene portato; per avere un mapping completo tra il Tipo di portare e i verbi in lingua cinese si devono però inserire nuove scene in cui questo tratto (il modo), che è irrilevante per *portare*, viene modificato, ottenendo così azioni in cui si porta l'oggetto in braccio, sulle spalle, in mano, ecc.. In questo caso quindi le nuove scene non sono istanze del prototipo di *portare*, ma altri prototipi, che si collocano allo stesso livello. Questo esempio evidenzia quindi che il vincolo di un solo prototipo per Tipo, che si ha nella struttura dati gerarchica, è troppo rigido per un'estensione di IMAGACT in un contesto interlinguistico: il modello logico deve consentire l'inserimento di più prototipi azionali per uno stesso Tipo.

Per poter soddisfare il requisito di arricchimento dell'ontologia con nuove scene, la struttura dati deve quindi:

- consentire l'inserimento di scene che siano istanze di un prototipo azionale;
- permettere di aggiungere altre scene prototipiche ad un Tipo.

A questi due punti nodali si aggiungono dei requisiti minori che la struttura dati deve garantire: uno di questi è la possibilità di inserire perifrasi per fornire una traduzione in quei casi in cui il riferimento all'azione avviene attraverso strutture non verbali (caso 4). Deve essere poi possibile lasciare una Scena senza collegamento a verbi, per poter gestire i casi in cui in una lingua manca il concetto azionale (caso 6); questo è un requisito abbastanza banale per la struttura dati, mentre è un punto da tenere in considerazione nello sfruttamento dell'ontologia.

RISTRUTTURAZIONE DELL'ONTOLOGIA

I casi di studio presentati hanno evidenziato una serie di limiti strutturali di IMAGACT che portano ad una generale difficoltà di gestione dell'ontologia, in cui operazioni di base come inserimenti, cancellazioni e modifiche sui dati non sono sempre sicure e possono inficiare la coerenza dell'informazione. Le seguenti operazioni devono essere possibili nel nuovo modello dati:

1. rappresentazione dei rapporti tra tipi attraverso intersezione insiemistica;
2. inserimento di una Scena che sia il riferimento di due tipi semanticamente diversi;
3. rimozione o inserimento di una Scena da un Tipo senza modificarne le proprietà;
4. definizione dei tipi in modo indipendente dagli altri;
5. rappresentazione dell'equivalenza locale di due verbi senza che questa implichi una relazione semantica tra essi;
6. inserimento di scene che siano istanze di un prototipo azionale;
7. aggiunta di altre scene prototipiche ad un Tipo.

Soltanto l'opzione (6) è possibile nel modello gerarchico, mentre la (4) non è sempre garantita e tutte le altre operazioni non sono mai possibili.

6.1 IL MODELLO INSIEMISTICO

A partire da queste considerazioni è stato sviluppato un nuovo modello dati per IMAGACT, in grado di superare i limiti del modello gerarchico e di soddisfare i requisiti funzionali.

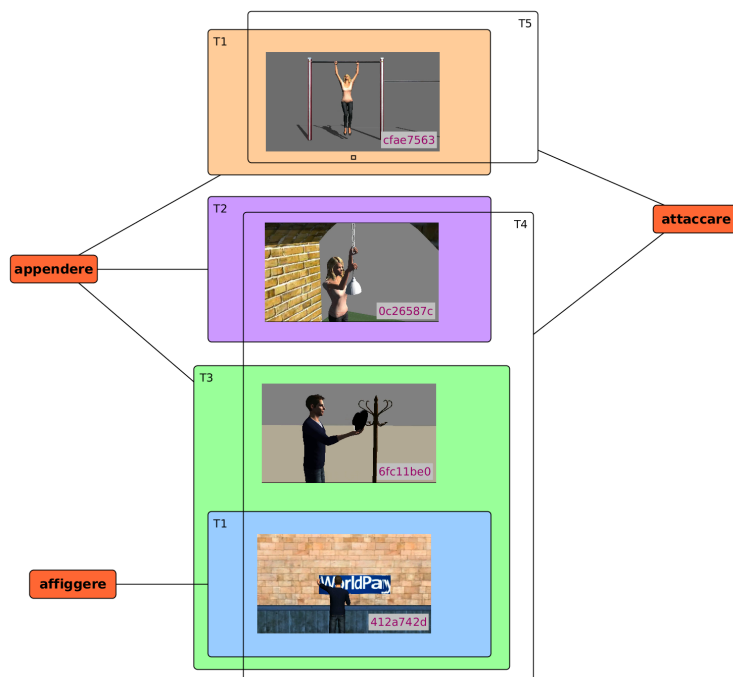
6.1.1 *La semplificazione dei concetti di IMAGACT*

La nuova struttura dati semplifica la relazione tra i concetti azionali e le scene prototipiche, eliminando quell'ambiguità degli oggetti Scena descritta nel paragrafo 4.3.1: nella nuova rappresentazione, infatti, il Tipo è l'unica entità demandata alla categorizzazione dei concetti azionali, mentre la Scena rappresenta semplicemente un prototipo azionale, come peraltro era previsto nel modello semantico originale di IMAGACT. Questa ridefinizione degli oggetti Scena risulta imprescindibile, in quanto diretta conseguenza dei requisiti funzionali, in particolare di (2) e (3). In tal modo la Scena, perdendo il suo ruolo di categoria azionale definita sulla base di tratti linguistici inerenti, torna ad essere un oggetto completamente indipendente dalla lingua, in grado di mettere in connessione i concetti azionali di lingue diverse. Questi concetti sono rappresentati dai Tipi, per cui la necessità di inserire forzature di categorizzazione per mezzo di scene *ad hoc*, come in 5.1.1, viene meno. In questa nuova struttura i Tipi preservano il loro ruolo di rappresentanti di concetti azionali: essi si riferiscono ad uno e un solo verbo e ne partizionano la variazione verticale. Ciò che cambia nei Tipi è, ovviamente, il loro rapporto con il riferimento. Essi infatti divengono dei contenitori di Scene: all'interno di un Tipo azionale vengono racchiuse tutte le scene che lo rappresentano prototipicamente. Questo costituisce una soluzione per (3) e (7), poiché permette ai tipi di avere più di un prototipo e di modificarne il riferimento senza alterarne la semantica. Da questo punto di vista il Tipo può essere identificato esplicitamente da una "famiglia" anche nel modello ontologico e non solo dal punto di vista della visualizzazione dei dati.

L'esempio di figura 49 mostra come risulti cambiata la rappresentazione dei tre verbi *attaccare*, *appendere* e *affiggere* nel modello insiemistico (si rimanda alla Fig. 21 per lo stesso esempio nel modello gerarchico); per chiarezza sono stati riportati soltanto due tipi di *attaccare*. Qui *affiggere* ha un solo Tipo, che contiene una sola Scena ("412a742d"), la quale però fa parte anche di *appendere-T3* e di *attaccare-T4*. Ne segue che *attaccare*, *appendere* e *affiggere* sono localmente equivalenti nello specifico campo di variazione rappresentato da "412a742d". Notiamo inoltre come *attaccare* sia sempre applicabile nell'intera variazione di *appendere*: infatti le scene contenute in tutti e tre i tipi di *appendere* fanno parte anche dei tipi di *attaccare*. Inoltre, mentre negli altri campi di variazione *appendere* è più specifico

di *attaccare* (*appendere-T2* e *appendere-T3*), il primo Tipo di *appendere* e *attaccare-T5* sono equivalenti.

Figura 49: La rappresentazione dei rapporti tra *affiggere*, *appendere* e *attaccare* nel modello insiemistico



La conseguenza teorica più forte che questo nuovo modello porta con sé è il fatto che i Tipi azionali siano dipendenti dalla lingua, e più specificamente dal verbo che si prende in esame. Con lo sviluppo dell'ontologia, infatti, sono emersi due fattori determinanti per la rimodellazione della base dati. Il primo di questi consiste nell'evidenza che i Tipi azionali possono avere livelli diversi di generalità (come anche rappresentato attraverso la relazione di istanza). Il secondo aspetto coincide con l'osservazione che i concetti azionali sono discretizzati in modo diverso dai singoli verbi nelle varie lingue: in sintesi, non tutti i concetti ipotizzabili a partire da una lingua sono riscontrabili in tutte le altre. Quest'ultimo aspetto è emerso principalmente in connessione allo studio di lingue tipologicamente diverse dall'italiano e dall'inglese, e in particolare del cinese. In realtà anche tra due lingue molto lontane (come l'italiano e il cinese), la condivisione dei concetti è mediamente molto alta (84,14%, vedi Par. 8.1). Allo stesso tempo, però, non è detto che un concetto identificato in alcune lingue sia valido anche per le altre: attraverso l'analisi

linguistica, non siamo in grado di definire in modo categorico concetti azionali universali. La suddivisione del campo semantico dell'azione avviene per mezzo di differenziali linguistici, che possono però essere annullati in alcune lingue. Per questa ragione, avere la possibilità di rappresentare più liberamente i diversi casi di categorizzazione linguistica, senza necessariamente ancorarli a categorie fisse, è necessario soprattutto a garantire l'estendibilità dell'ontologia a lingue finora non trattate.

6.1.2 *Un nuovo modello inferenziale*

Se da un lato il modello gerarchico di IMAGACT non è sufficientemente espressivo per una rappresentazione coerente dell'informazione e limita le possibilità di sviluppo dell'ontologia, dall'altro la strutturazione rigida dei dati si presta bene alla formulazione di inferenze logiche tra verbi e tipi. In questo nuovo modello invece le entità sono più fluide e i rapporti tra di esse meno strutturati, pertanto è stato necessario imporre dei vincoli all'ontologia, che altrimenti sarebbe risultata troppo libera e con scarso potere inferenziale. In particolare i requisiti (2), (4) e (5) portano ad un modello in cui i Tipi sono tra loro semanticamente indipendenti ed è quindi necessario esplicitare in che modo si esprimono le dipendenze e, in generale, tutte le relazioni logiche che nel modello gerarchico venivano determinate dall'incapsulamento dei dati.

Di seguito si riporta la riformulazione delle inferenze descritte nel modello gerarchico:

- **Inferenza diretta tra verbi**

Dati due verbi v_1 e v_2 , l'enunciato "se v_1 allora v_2 " è vero se e solo se tutte le scene di tutti i Tipi di v_1 sono anche scene di uno o più tipi di v_2 ;

- **Inferenza generale tra Tipo e verbo**

Dato un Tipo t_1 e un verbo v_2 , l'enunciato "se t_1 allora v_2 " è vero se per ogni Scena s di t_1 , s appartiene anche a un Tipo di v_2 ;

- **Applicabilità di un verbo ad una Scena**

Dati una Scena s_1 e un verbo v_1 , v_1 è applicabile a s_1 se e solo se esiste un Tipo t_1 di v_1 tale che t_1 contiene s_1 ;

- **Equivalenza locale di due verbi**

Dati due verbi v_1 e v_2 , v_1 è localmente equivalente a v_2 se e solo se esistono un Tipo t_1 di v_1 e un Tipo t_2 di v_2 tali che t_1 e t_2 hanno almeno una Scena in comune;

- **Traducibilità di due verbi in lingue diverse**

Dati un verbo v_1 in una lingua L_1 e un verbo v_2 in una lingua L_2 diversa da L_1 , v_2 è un traduce di v_1 se e solo se v_1 e v_2 sono localmente equivalenti.

In queste nuove definizioni c'è una evidente semplificazione, derivata dall'utilizzo del modello insiemistico, più sintetico ed espressivo di quello gerarchico; oltre a ciò però emerge anche una differenza cruciale nella concezione dell'equivalenza locale: se prima infatti era il Tipo azionale a determinare l'ambito in cui due verbi potevano equivalersi, adesso è possibile avere equivalenza locale anche per concetti azionali più specifici rispetto a quelli definiti da un Tipo. Questo accade in tutti quei casi in cui si ha intersezione insiemistica tra tipi, come nell'esempio di Fig. 37 o nella figura seguente 50, dove *bucare* è localmente equivalente a *passare* nonostante che l'equivalenza si realizzi soltanto in una parte dell'estensione dei tipi. Si nota inoltre che non possiamo formulare implicazioni tra Tipo e verbo, ma abbiamo comunque equivalenza locale. Questo demarca una notevole differenza rispetto alla struttura dati gerarchica in cui il Tipo definiva il campo di variazione entro cui i verbi erano equivalenti e quindi potevano sempre essere formulate inferenze tra Tipo e verbo in tutti i casi di equivalenza locale.

Figura 50: L'intersezione tra due concetti azionali legati ai verbi *bucare* e *passare*



Nel modello gerarchico l'equivalenza locale era definita di fatto sulla base di un Tipo:

Due verbi sono localmente equivalenti quando entrambi si applicano alle istanze di un concetto azionale rappresentato da un Tipo

Questa definizione esplicita chiaramente cosa significhi “localmente”, ossia il campo di variazione definito da un Tipo azionale. La nuova definizione di equivalenza locale, invece, può essere così formulata:

Due verbi sono localmente equivalenti quando condividono un prototipo azionale

In questa riformulazione non è più così chiaro cosa significhi “localmente”: considerando l’esempio di figura 50, sappiamo per certo che (a) i due verbi non sono localmente equivalenti nel concetto rappresentato da *bucare-T1*, (b) non sono localmente equivalenti nel concetto rappresentato da *passare-T6*, (c) i due verbi sono entrambi applicabili all’azione specifica rappresentata dalla Scena prototipo “7e1bc593”. Deriviamo quindi che o l’equivalenza locale tra questi due verbi è limitata ad una sola azione oppure è circoscritta ad una classe di azioni simili a “7e1bc593”, ma diverse sia da “906e648b” che da “4ce0740f”. Possiamo scartare la prima ipotesi trovando un semplice controesempio, come “L’operaio buca la montagna con la trivella / La trivella passa la montagna”, che è un’azione diversa da “7e1bc593”, in cui però vale l’equivalenza locale. L’evidenza di produttività ci dice quindi che deve esistere un concetto azionale più specifico di quelli rappresentati dai due Tipi di *passare* e *bucare*, del quale la Scena “7e1bc593” rappresenta un prototipo. Se così non fosse non potremmo avere produttività e quindi applicare i due verbi equivalenti ad azioni simili a “7e1bc593”. Questo nuovo concetto che emerge dall’intersezione di due tipi *bucare-T1* e *passare-T6* potrebbe intuitivamente corrispondere ad un Tipo di un verbo italiano non presente in IMAGACT, come ad esempio *traforare* o *trapassare*; in generale però non possiamo fare alcuna assunzione sul fatto che quel concetto sia lessicalizzato in una lingua. Ad esempio, nel caso di *scaricare-svuotare* (Par. 5.1.1) è già più difficile trovare un corrispettivo verbale che unisca questi due concetti.

In questo quadro teorico abbiamo quindi nuovi concetti azionali che derivano da intersezioni di tipi, ma che non hanno (necessariamente) una lessicalizzazione, ossia non corrispondono a nessun Tipo. Questi casi, che intra-linguisticamente sono abbastanza rari (almeno per le evidenze che si hanno dall’italiano e dall’inglese) acquistano un grande peso in contesto interlinguistico, dove la sovrapposizione tra tipi è diffusa, dal

momento che la categorizzazione dello spazio azionale differisce tra lingua e lingua. Possiamo quindi immaginare che il numero di nuovi concetti derivanti da intersezioni di tipi aumenterà molto via via che verranno mappate nuove lingue in IMAGACT. In una ipotetica ontologia completa, in cui cioè tutti i verbi d'azione delle lingue rappresentate sono categorizzati (suddivisi in tipi azionali) e mappati, un concetto azionale può essere:

- esteso a tutta l'ontologia: ogni lingua dell'ontologia ha almeno un Tipo che rappresenta quel concetto;
- parzialmente condiviso: due o più lingue dell'ontologia, ma non tutte, hanno almeno un Tipo che rappresenta quel concetto;
- dipendente dalla lingua: solo una lingua ha un Tipo per quel concetto;
- non linguistico: nessuna lingua ha tipi per quel concetto.

Come abbiamo detto, per permettere di effettuare inferenze corrette nel nuovo modello, è necessario imporre dei vincoli, che sono già stati implicitamente espressi dalle definizioni di inferenze tipo-verbo e verbo-verbo. La prima restrizione, che ci consente di rappresentare l'equivalenza di due tipi, è la seguente:

Se due tipi contengono lo stesso set di scene, allora rappresentano lo stesso concetto azionale

Notiamo che, in linea di principio, due tipi potrebbero essere definiti su concetti semantici indipendenti, ma avere le stesse scene come riferimento prototipico. Se si considera l'esempio 5.1.1 (*scaricare-svuotare*), la Scena "do69151b" è un buon prototipo sia per *scaricare* che per *svuotare*; per quanto detto sopra, se i due tipi non hanno altre scene prototipo, è possibile inferire che quei due tipi sono equivalenti e conseguentemente formulare inferenze come "se scarico-T₁ allora svuoto" e "se svuoto-T₁ allora scarico". Naturalmente ciò non è corretto, ma anzi è profondamente inconsistente con la semantica dei due verbi in questione e la prova di questo è data dal fatto che sia possibile trovare una serie di altre azioni appartenenti allo stesso Tipo di *svuotare*, ma che non sono scaricare (ad esempio "Marco svuota la credenza") e viceversa. Proprio sfruttando questa proprietà possiamo rappresentare i tipi in modo differenziale, aggiungendo altre scene

che appartengono allo stesso Tipo di *svuotare* ma non sono *scaricare* e che appartengono allo stesso Tipo di *scaricare* ma non sono *svuotare*. In tal modo si crea un'intersezione tra tipi, la quale determina equivalenza locale, ma non relazione semantica tra concetti linguisticamente individuati.

L'aggiunta di scene differenziali è quindi la prassi metodologica corretta per poter mappare tra loro due tipi che fanno riferimento a concetti diversi. Parallelamente il vincolo esposto sopra ci consente di rappresentare facilmente l'equiestensionalità e quindi di formulare inferenze tipo-verbo.

Il secondo vincolo è relativo alla rappresentazione di relazioni gerarchiche tra tipi:

Se un Tipo t_1 contiene un sottoinsieme delle scene contenute in un altro Tipo t_2 allora il concetto azionale espresso da t_1 è più specifico di quello espresso da t_2

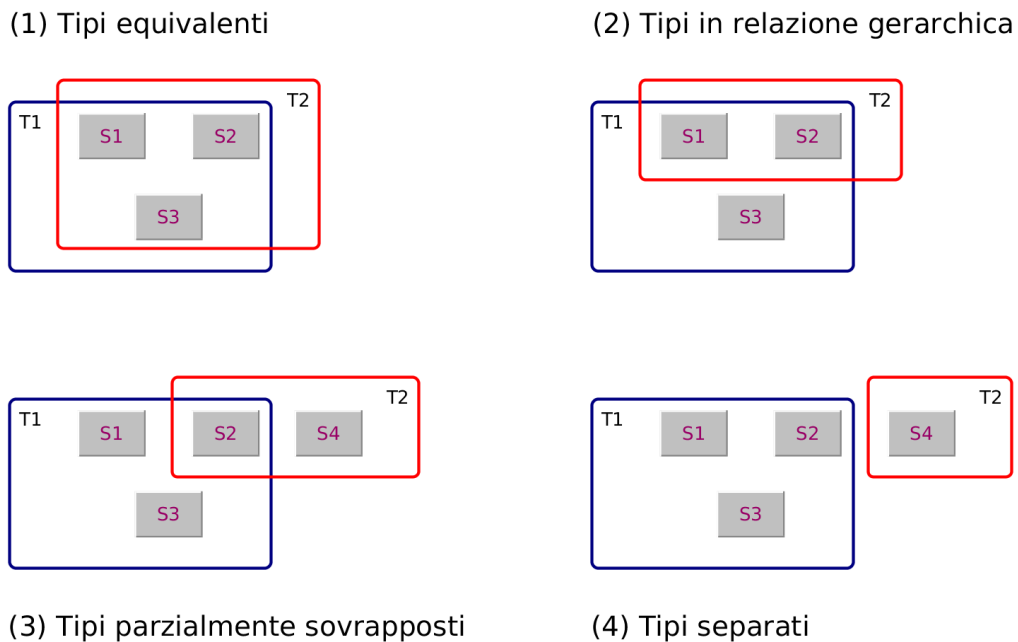
In questo caso è possibile inferire che t_1 implica il verbo di t_2 , ma non il viceversa, poiché t_2 rappresenta un concetto più generale. Similmente a quanto accadeva per l'equivalenza, questo vincolo determina la necessità di inserire nuove scene differenziali in un Tipo se è necessario specificarne l'indipendenza semantica da un altro con cui esso condivide dei prototipi.

Riassumendo, due tipi tra loro possono essere:

- equivalenti, se contengono le stesse scene (1): $t_1 \cap t_2 = t_1 = t_2$
- in relazione gerarchica, se uno contiene un sottoinsieme delle scene dell'altro (2):
 $t_1 \cap t_2 = t_2$
- indipendenti ma condividere un prototipo, se la loro intersezione non è vuota (3):
 $t_1 \neq t_1 \cap t_2 \neq t_2$
- separati se la loro intersezione è vuota (4): $t_1 \cap t_2 = \emptyset$

Sofferamoci ancora sul caso (3) in cui due Tipi sono parzialmente sovrapposti, che è il punto in cui emerge la maggiore espressività del modello insiemistico rispetto a quello gerarchico. Come abbiamo detto, in questi casi i due concetti azionali sono definiti indipendentemente, ma condividono alcuni prototipi pragmatici. Questo non significa che i due Tipi siano tra loro completamente separati come invece accade in (4), ma

Figura 51: Relazioni possibili tra Tipi



invece essi risultano tra loro pragmaticamente relati in relazione all'ampiezza della loro intersezione. Banalmente non possono essere trattati allo stesso modo due Tipi in cui tutti i prototipi sono differenziali ad eccezione di uno che è comune ad entrambi e due Tipi che condividono la maggior parte dei prototipi ma risultano differenziati da poche scene isolate. La cardinalità dell'intersezione può quindi essere utilizzata per misurare la vicinanza tra due Tipi che sono in relazione di sovrapposizione parziale. Dati due Tipi t_1 e t_2 possiamo misurare

$$sim(t_1, t_2) = \frac{|t_1 \cap t_2|}{|t_1 \cup t_2|}$$

In tal modo laddove non possiamo dire che due Tipi sono equivalenti (1), in relazione gerarchica (2) o separati (4), possiamo fornire un valore che ne misura la vicinanza categoriale. Questo è significativo in un contesto interlinguistico dove il numero di Tipi parzialmente sovrapposti (3) è destinato a crescere. Parallelamente anche il numero di scene crescerà via via che il mapping di nuove lingue porterà nell'ontologia altri concetti azionali e, conseguentemente, aumenterà la precisione della misura *sim*.

6.2 L'ADEGUATEZZA DEL MODELLO

Mostriamo adesso come il nuovo modello dati sia in grado di soddisfare i requisiti funzionali espressi all'inizio del capitolo. Il primo, cioè la capacità di rappresentare l'intersezione insiemistica tra tipi, costituisce il salto espressivo della nuova struttura dati rispetto al modello gerarchico ed è stato già ampiamente descritto. Il secondo requisito è relativo alla possibilità che una stessa Scena possa fungere da prototipo per due tipi che rappresentano concetti azionali diversi. Come abbiamo visto nel paragrafo precedente, questo è possibile se i due tipi sono differenziati da altre scene: infatti è necessario che l'intersezione tra i due tipi sia diversa da ciascuno di essi. Quando ciò non accade, i vincoli definiti sul modello dati portano ad inferire una relazione semantica tra i tipi (di dipendenza gerarchica o di equivalenza). Pertanto il requisito è soddisfatto, purché oltre alla Scena in questione si inseriscano tutte le altre scene necessarie a garantire la rappresentazione differenziale dei tipi. Similmente, anche l'aggiunta e la rimozione di scene devono tenere conto del modello inferenziale. Di seguito si riporta la casistica possibile, facendo riferimento alle generiche relazioni tra tipi (identificati come t_1 e t_2) rappresentate in Fig. 51:

- se t_1 e t_2 sono equivalenti, l'aggiunta di una Scena a t_1 e non a t_2 mette i tipi in relazione gerarchica con t_1 più generale di t_2 ;
- se t_1 e t_2 sono equivalenti, la rimozione di una Scena da t_1 e non da t_2 mette i tipi in relazione gerarchica con t_2 più generale di t_1 ;
- se t_1 e t_2 sono in relazione gerarchica con t_1 più generale di t_2 , l'aggiunta di una Scena a t_1 e non a t_2 non ha conseguenze particolari;
- se t_1 e t_2 sono in relazione gerarchica con t_1 più generale di t_2 , l'aggiunta di una Scena a t_2 e non a t_1 crea un'intersezione tra concetti;
- se t_1 e t_2 sono in relazione gerarchica con t_1 più generale di t_2 , la rimozione di una Scena da t_1 e non da t_2
 - non ha conseguenze, se t_1 contiene almeno un'altra Scena che non fa parte di t_2 ;

- crea due tipi equivalenti, altrimenti.
- se t_1 e t_2 sono in relazione gerarchica con t_1 più generale di t_2 , la rimozione di una Scena da t_2 e non da t_1
 - non ha conseguenze, se t_2 contiene almeno un'altra Scena;
 - provoca l'eliminazione di t_2 altrimenti.
- se t_1 e t_2 sono parzialmente sovrapposti, l'aggiunta di una Scena a t_1 e non a t_2 non ha conseguenze (e viceversa);
- se t_1 e t_2 sono parzialmente sovrapposti, la rimozione di una Scena da t_1 e non da t_2
 - crea due tipi separati, se la Scena era l'unica intersezione tra t_1 e t_2 ;
 - crea una relazione gerarchica con t_2 più generale di t_1 , se la Scena non era parte dell'intersezione e t_1 non ha altre scene fuori dall'intersezione;
 - non ha conseguenze, altrimenti.

Sebbene questa classificazione debba essere presa in considerazione per le operazioni di inserimento e rimozione di scene su tipi, bisogna tenere presente che il *framework* di IMAGACT si fonda su dati empirici e non su definizioni di concetti a priori. Di conseguenza i rapporti tra tipi azionali non sono predefiniti, ma derivano dalla loro possibilità di applicarsi alle azioni fisiche. In questo quadro il mapping tra le varie lingue non è più un'operazione fatta dall'annotatore (come invece è avvenuto nella fase di creazione dell'ontologia), ma è derivata automaticamente dai dati di annotazione.

Il quarto requisito si riferisce alla possibilità di definire i tipi in modo indipendente dagli altri: questo è garantito nel modello, in cui non solo i tipi hanno una definizione indipendente, ma, utilizzando le scene differenziali, si possono legare anche a scene di altri tipi, preservando l'indipendenza semantica da essi. Se due tipi di due verbi diversi condividono un prototipo azionale, allora quei verbi sono localmente equivalenti. Come abbiamo visto però questo non comporta che i due verbi debbano avere una relazione semantica: in particolare, questo accade nei casi di intersezione tra tipi.

Infine gli ultimi due requisiti, relativi alla possibilità di inserire nuove scene istanza di un Tipo e nuove scene prototipo, sono garantiti dal modello dati e dai vincoli di

inferenza. La struttura infatti permette di avere più di un prototipo per ogni Tipo e i vincoli imposti (Par. 6.1.2) consentono di formulare inferenze tra tipi più specifici e più generali utilizzando l'insieme di scene comuni.

6.2.1 *Esempi sui casi di studio*

Per mostrare l'adeguatezza del modello vengono qui presi in esame alcuni casi di studio dettagliati nel capitolo precedente: tali esempi infatti rappresentano le classi di problemi che sono emersi dall'analisi del modello dati originale. In particolare vengono descritte le operazioni che sono necessarie per fornire una rappresentazione coerente dell'informazione nella nuova struttura dati insiemistica.

ESEMPIO 1: SVUOTARE E SCARICARE La relazione tra svuotare e scaricare non è gerarchica: questi due lemmi infatti si estendono alla stessa classe di azioni, che possiamo descrivere come "rimuovere completamente il contenuto solido da un veicolo", ma che di per sé non costituisce un campo di variazione verticale né per scaricare, né per svuotare. Per entrambi i verbi, infatti, questa classe di azioni è troppo ristretta per essere un Tipo; essa cioè contiene troppi tratti semantici. Per il Tipo di scaricare il tratto determinato dal modificatore "completamente" è irrilevante: il fatto che gli oggetti vengano rimossi dal veicolo siano tutti o soltanto una parte non determina un diverso concetto azionale. Per il Tipo di svuotare il fatto che il tema sia un "veicolo" è irrilevante: modificazioni dell'azione con altri oggetti che fungono da contenitori (come un mobile o una scatola) non portano a variazione tipologica. Il modello insiemistico è sufficientemente espressivo per rappresentare questa situazione di sovrapposizione parziale tra tipi azionali. In particolare sono necessarie tre scene, (a) una in cui si scarica e non si svuota, (b) una in cui si svuota e non si scarica e (c) una in cui si svuota e si scarica.

Il Tipo di scaricare è creato come l'insieme delle scene (a) e (c); il Tipo di svuotare è creato come l'insieme delle scene (b) e (c). In tal modo si formalizza correttamente il rapporto tra i due tipi che non hanno una specifica relazione semantica, ma sono entrambi applicabili ad un contesto pragmatico.

Figura 52: Scene necessarie per rappresentare il differenziale tra *svuotare* e *scaricare*

(a) scaricare

(b) svuotare

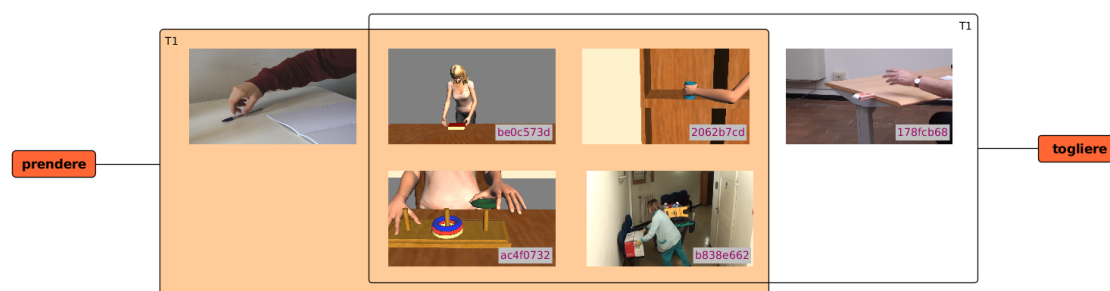
(c) scaricare e svuotare

ESEMPIO 2: PRENDERE E TOGLIERE Il rapporto tra i verbi prendere e togliere si esprime in modo analogo al caso precedente: infatti, benché i due tipi di questi verbi sembrino essere in relazione gerarchica l'uno con l'altro (*se prendo allora tolgo*), esistono casi in cui prendere non è predicabile con togliere, ad esempio in una frase come "Marco prende la penna". Pertanto anche in questo caso la situazione che il modello insiemistico deve rappresentare è quella di una sovrapposizione parziale di due tipi, che condividono alcuni prototipi, ma non tutti. La correzione della struttura dati prevede in questo caso:

- la rimozione della Scena duplicata di Fig. 39 (*Marta prende il bicchiere dal tavolo*) ed il collegamento del Tipo di *togliere* alla Scena di *prendere* (*Fabio prende la tazza dal ripiano*, Fig. 38);
- la creazione di due scene differenziali, una in cui si prende e non si toglie e una in cui si toglie e non si prende.

Una volta fatto questo, la relazione tra i tipi di *prendere* e di *togliere* è rappresentata in modo coerente e differenziale. Vediamo quindi che la struttura minima per formalizzare un caso di sovrapposizione parziale dei concetti prevede tre scene. A differenza del caso precedente, i verbi *prendere* e *togliere* sono molto generali ed anche i loro Tipi categorizzano una classe ampia di azioni; pertanto ciascuno di questi tipi contiene più scene. Questo non costituisce un limite nella struttura dati poiché sopra la struttura minimale possono essere aggiunte liberamente nuove scene ad un Tipo o ad un altro (o a tutti e due).

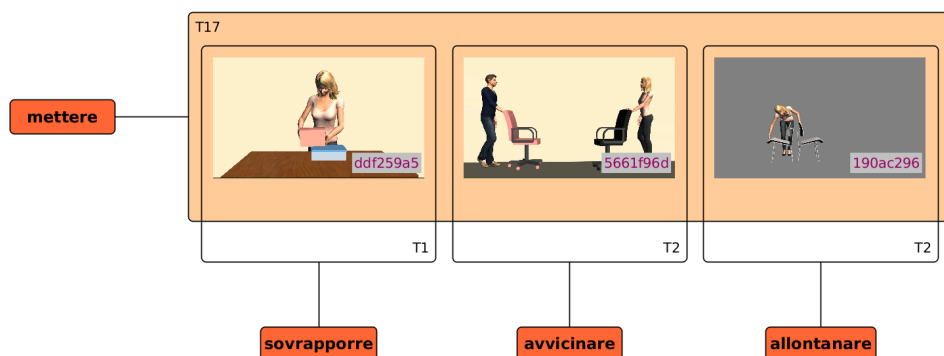
ESEMPIO 3: METTERE Supponendo che la Scena di figura 41 (*Fabio e Marco mettono le due sedie più vicine*) rappresenti un concetto pragmatico esterno alla variazione orizzontale del primo Tipo di *mettere*, nel nuovo modello dati è sufficiente rimuovere quella Scena

Figura 53: Formalizzazione del rapporto tra *prendere* e *togliere*

dal Tipo ed inserirla in un altro. Questa modifica infatti non si ripercuote sugli altri tipi (equivalenti, più generali o più specifici) che contengono quella Scena, ma semplicemente cambia la loro relazione con il Tipo di *mettere*, secondo le regole viste all'inizio di questo paragrafo.

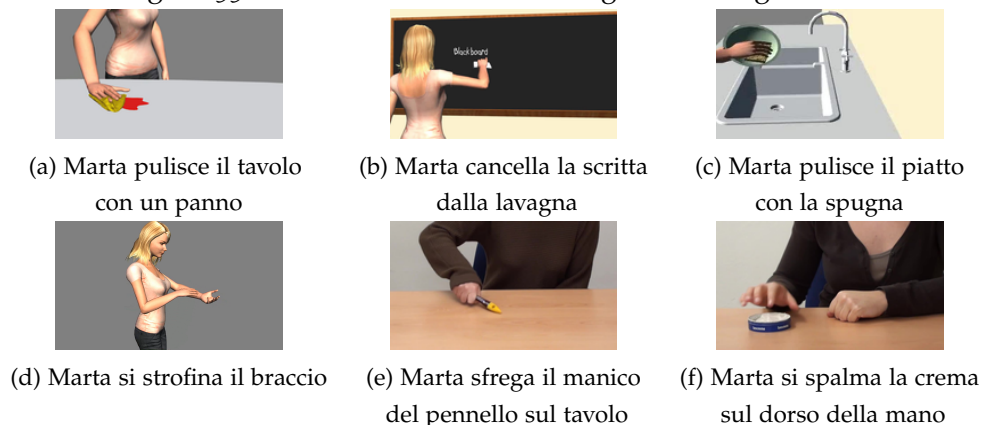
Ipotizziamo, ad esempio, che un revisore trovi la necessità di creare un nuovo Tipo di *mettere* (*mettere-T17*) concepito sul tratto del “cambiamento della relazione spaziale tra due oggetti” (*mettere due oggetti vicini / lontani / attaccati / sovrapposti / ...*), che nella versione attuale non è stato considerato discriminativo; difatti tutti i concetti azionali di questa classe sono confluiti nel Tipo molto generale *mettere-T1*, concepito sul tratto generico di “collocazione di un oggetto in un luogo” (Fig. 40). Per effettuare la modifica suddetta si crea un nuovo Tipo (*mettere-T17*) che racchiude tutte le scene di *mettere-T1* nelle quali si cambia una relazione spaziale tra due oggetti e si tolgono quelle scene da *mettere-T1*. Queste operazioni non hanno effetto sugli altri verbi della gerarchia, i quali rimangono collegati alle stesse scene. Nella figura seguente è mostrato il nuovo Tipo *mettere-T17* dopo le modifiche descritte; per semplicità sono state riportate soltanto tre scene e quattro verbi equivalenti.

ESEMPIO 4: PULIRE Il caso relativo alla rappresentazione dei verbi *pulire* e *suonare* ha generato un problema di mapping con la lingua cinese in cui questi concetti generali non sono lessicalizzati in una struttura verbale, la quale è invece demandata a descrivere l'azione specifica che si compie quando si pulisce o quando si suona. Questo fatto non costituisce di per sé un problema (lo si può annoverare tra i casi di *lexical gap*), ma lo diventa quando, all'interno della struttura dati gerarchica, volessimo creare i tipi

Figura 54: Ipotesi di creazione di un nuovo Tipo di *mettere*

azionali per la lingua cinese. Infatti saremmo costretti ad eseguire una discriminazione dei concetti azionali non soltanto sulla base della variazione primaria del lemma cinese, ma tenendo conto anche della gerarchia di concetti italiano-inglese esistente. Ad esempio il verbo 擦 (*cā*), che predica di un'azione eseguita per sfregamento, si riferisce ad 11 concetti azionali distinti, alcuni dei quali rientrano nella variazione di pulire (Fig. 55, esempi (a), (b) e (c)), mentre altri no (esempi (d) (e) e (f)).

Figura 55: Serie di azioni che coinvolgono lo sfregamento



All'interno dell'estensione di 擦 (*cā*) soltanto un madrelingua cinese è in grado di giudicare quali sono i nuclei di variazione primaria: ad esempio la presenza o meno del tratto di "rendere pulita una superficie" potrebbe essere un parametro importante nella semantica di 擦 (*cā*), che si basa sull'idea di "sfregamento", e quindi determinare una suddivisione tipologica, oppure potrebbe essere totalmente irrilevante. In questo secondo caso 擦 (*cā*) non sarebbe mappabile nel modello gerarchico di IMAGACT. Viceversa per

la proprietà di indipendenza dei tipi che è alla base del modello insiemistico, la nuova struttura è in grado di consentire una qualsiasi suddivisione in Tipi, raggruppando i video sulla base della similarità (pragmatica e semantica).

IMPLEMENTAZIONE IN UN DATABASE A GRAFO

Accanto alle questioni di modellazione logica dei dati discussi precedentemente, che hanno portato ad una necessità di ristrutturazione del database sulla base di un nuovo modello teorico, esistono anche delle problematiche relative alla codifica dei dati in un modello relazionale. Alcuni limiti dei database relazionali infatti sono emersi come punti critici per la rappresentazione e l'estensione dell'ontologia:

- rigidità strutturale: la rigidità del modello relazionale può rappresentare un problema nello sviluppo di una risorsa linguistica di questo Tipo dove i livelli di annotazione sono molti e l'informazione viene estesa in modo incrementale. L'inserimento di nuovi livelli di annotazione, che spesso è circoscritta ad un numero limitato di oggetti, porta ad avere tabelle con *sparse data*, cioè con molte colonne e molti valori *NULL*.
- distanza dal modello insiemistico: un modello dati insiemistico si può rappresentare bene in un database relazionale, ma non in modo diretto per cui abbiamo un certo grado di *impedance mismatch*. Questo costituisce un problema abbastanza rilevante per IMAGACT poiché il modello dati astratto è destinato ad evolversi e il costo delle modifiche deve essere ridotto al minimo.
- distanza da un'ontologia formale RDF: una proprietà importante del nuovo database è la sua vicinanza ad un modello ontologico RDF, in modo che la conversione tra i dati sia diretta. La struttura logica di un database relazionale è molto lontana da un'ontologia e questo rende necessario avere due strutture dati diverse (una tabellare e una sotto forma di triple); pertanto ogni modifica alla struttura del database dovrebbe essere riportata anche sul modello ontologico attraverso una procedura di conversione.

Queste ragioni ci hanno indotto ad esplorare la classe dei database NoSQL ed in particolare i database a grafo, che, come vedremo, sono risultati più adeguati a soddisfare i requisiti strutturali e funzionali di IMAGACT.

7.1 NOSQL DATABASE

Il database relazionale costituisce oggi il sistema di *data persistence* più utilizzato al mondo, grazie alle sue caratteristiche di affidabilità, sicurezza, espressività, integrabilità. Inoltre gli RDBMS garantiscono facilità di accesso ai dati attraverso il linguaggio quasi-standard SQL, alte prestazioni, gestione della concorrenza, API e librerie in un'ampia varietà di linguaggi. Tutto ciò ha fatto sì che i database relazionali siano stati protagonisti del mercato per oltre 30 anni. Negli ultimi anni però nuovi strumenti software che rientrano sotto il nome di "database NoSQL" (acronimo di Not Only SQL) hanno fatto emergere una varietà di tecnologie per la rappresentazione persistente dei dati diverse dal modello relazionale.

I database NoSQL appartengono ad una categoria molto generale determinata soltanto in negativo dall'assenza di una proprietà, l'essere cioè diversi da un "database SQL", ossia un database relazionale interrogabile attraverso query SQL. Nonostante che questo renda molto estesa e varia la classe di software che rientrano sotto il nome di database NoSQL, sono stati evidenziati dei caratteri comuni a gran parte di questi strumenti (Huang and Luo, 2014):

- assenza di schema: i dati nel db possono essere inseriti in assenza di una struttura dati predefinita;
- architettura distribuita *fault tolerant*: il database è scalabile su un sistema distribuito e tollerante ai fallimenti hardware;
- *caching* integrato: molti database NoSQL integrano una cache per utilizzare la memoria in modo efficiente ed accelerare le operazioni frequenti.

La ricerca di soluzioni di memorizzazione alternative è stata spinta dalla necessità di superare alcuni problemi di cui soffrono i sistemi basati sul modello relazionale

come rigidità strutturale, *impedance mismatch*, bassa scalabilità e ridotte performance per transazioni “leggere”. In un database relazionale non è possibile aggiungere, rimuovere o modificare un campo o una proprietà di un oggetto senza alterare l’intera classe di quegli oggetti; questo perché ogni dato deve essere tipato e incardinato nello schema di tabelle e relazioni definito. Questo Tipo di rigidità, se da un lato consente di preservare la coerenza logica dei dati, dall’altro è poco compatibile con applicazioni dinamiche in cui la struttura dei dati si evolve nel tempo, dal momento che rende complesso e costoso l’adeguamento del db.

Un altro problema sopracitato è quello dell’*impedance mismatch*: il modello relazionale è molto espressivo e consente di rappresentare strutture dati variegata, ma non sempre in modo “leggibile”. Spesso infatti l’informazione è frammentata e un’unica entità logica è memorizzata su più tabelle, rendendo necessario riaggregare i dati per poterli sfruttare. In particolare maggiore è la distanza tra il modello logico astratto e il modello relazionale e maggiore è la difficoltà di gestione dei dati, poiché la modifica ad un singolo elemento può tradursi in molte query di aggiornamento. I problemi di rigidità e *impedance mismatch*, che normalmente hanno un basso impatto nei sistemi statici, si ingigantiscono quando si devono trattare Big Data e una molteplicità di *light transaction* tipiche delle applicazioni *web-oriented*. Non è un caso che la ricerca nel campo dei database NoSQL (quello che viene chiamato *movimento NoSQL*) è stata promossa da grandi attori del web come Amazon, Google e Facebook¹.

Un’altra questione portata avanti dal movimento NoSQL è quella relativa alla costruzione di strumenti in grado di adattarsi ad ambienti distribuiti. La scalabilità orizzontale di un database, ossia la sua possibilità di estendersi su macchine diverse al crescere del sistema, è una caratteristica in forte contrasto con le transazioni ACID² che sono supportate dai RDBMS. Queste features che normalmente vengono fornite dai database relazionali e che garantiscono coerenza e sicurezza dei dati costituiscono però un ostacolo alla scalabilità e alle performance del sistema in ambiente distribuito. Questo è ben formalizzato nel CAP-Theorem³, che dimostra come un database distribuito sia

¹ Amazon Dynamo, Google BigTable e Facebook Cassandra.

² ACID è l’acronimo di Atomicity, Consistency, Isolation, Durability, 4 proprietà che garantiscono l’affidabilità delle transazioni in un database.

³ Il CAP-Theorem (formulato da Brewer nel 2000 (Brewer, 2000) come CAP-principle e formalizzato da Gilbert e Lynch (2012) dimostra che non è possibile avere scalabilità orizzontale mantenendo dati coerenti

incompatibile con le proprietà dei sistemi *ACID-compliant*. È importante però notare che non per tutte le applicazioni è necessario avere sistemi di persistenza che garantiscano rigidamente la qualità del dato e il buon esito delle transazioni: a volte, specialmente nell'ambito di grandi applicazioni web, come ad esempio i social network, è preferibile rilassare qualche vincolo in favore di maggiori dinamicità e velocità. Per questo molti database NoSQL implementano transazioni BASE⁴ (anziché ACID) e riescono così ad essere molto più veloci e facilmente integrabili in sistemi dinamici, rinunciando ad alcuni vincoli qualitativi sui dati.

7.1.1 Tipologie dei database NoSQL

La categorizzazione di Ben Scofield (2010) ci offre una panoramica sui modelli logici che sottostanno ai database NoSQL disponibili oggi:

- **Key-Value Stores;**
- **Column-Family Stores;**
- **Document Databases;**
- **Graph Databases.**

Ciascuno di questi modelli ha delle qualità che lo rendono più o meno adatto per i diversi ambiti applicativi; di seguito presentiamo una breve panoramica.

I più semplici database NoSQL sono quelli che fanno riferimento al modello **Key-Value Store**. Questi possono essere considerati come una versione minimale del modello relazionale: la struttura dati infatti è formata da una coppia chiave-valore, in cui la chiave è un ID e il valore è un dato grezzo non tipato. Il livello applicativo è totalmente responsabile della gestione dei valori. Per la sua semplicità questo modello è molto adeguato per alcune transazioni "leggere" come la memorizzazione dei dati di una sessione web, specialmente in un ambiente distribuito; questi strumenti sono infatti caratterizzati da alte prestazioni e forte scalabilità. Ad oggi sono disponibili diversi

(Coherence), certezza nella comunicazione (Availability) e tolleranza alla perdita di messaggi (Partition tolerance). Al massimo due di queste tre proprietà possono essere presenti in un sistema distribuito.

⁴ BASE è l'acronimo di Basically Available, Soft state, Eventual consistency.

prodotti open-source che rientrano nella categoria dei *key-value store*, come Redis, Riak o MemcacheDB, ciascuno con proprie funzionalità e capacità di memorizzazione.

Una versione più strutturata del Key-Value Store è il **Column-family Store**; questi strumenti memorizzano i dati per colonne invece che per righe come avviene nei DB relazionali. L'elemento base di rappresentazione sono le colonne, ciascuna delle quali contiene una tupla; la strutturazione dell'informazione avviene grazie a due elementi di più alto livello, le *super-column* – che raggruppano più colonne – e le *column-family* – che possono contenere un insieme di colonne e *super-column*⁵. Questi database permettono quindi di definire strutture dati più complesse rispetto ai *key-value store*, ma come questi ultimi offrono alte prestazioni e non vincolano i dati in uno schema rigido, poiché le *column-family* non devono rispettare un modello dati definito come invece avviene per le tabelle in un db relazionale. Le suddette proprietà rendono i *column-family store* particolarmente adatti per i big data, che possono essere caricati velocemente e modellati a posteriori. A discapito di questi vantaggi c'è però il fatto che questi database non consentono intersezioni di dati complesse che permettano di aggregare dati in modo comparabile con quello dei database relazionali. Fanno parte di questa categoria Cassandra e Hbase.

Un'altra estensione del *Key-Value Store* è il modello **document-oriented** dove, anziché di chiave-valore viene rappresentata la coppia chiave-documento. Questo porta ad una notevole differenza, poiché i documenti, a differenza dei valori, sono *inspectable*. Pertanto è possibile rappresentare e interrogare oggetti strutturati complessi. Anche in questo caso però la gestione della struttura dati è a totale carico dell'applicativo e il DB non opera nessun controllo sulla loro coerenza. Fanno parte di questa classe, ad esempio, i database XML che memorizzano file XML e possono essere interrogati attraverso Xpath o Xquery. Alcuni esempi noti di *document database* sono exist e MongoDB.

I database a grafo costituiscono un'alternativa ai database relazionali, poiché sono in grado di rappresentare dati strutturati di qualsiasi Tipo in un modello logico complesso. Sono pertanto adattabili a un'ampia varietà di domini diversi come i db relazionali, ma si basano sulla rappresentazione dell'informazione in termini di nodi e archi, anziché di tabelle e relazioni. Possono essere applicati in ambiti diversi, in tutti quei casi in cui la

⁵ Per questa descrizione è stato preso in considerazione il modello dati di Cassandra.

modellazione attraverso un grafo risulta più semplice e efficace di quella con le tabelle. Inoltre il database a grafo è più flessibile di un modello relazionale: scala con più facilità, il modello dati non è rigido e spesso l'aggregazione di informazioni si traduce in query molto meno complesse.

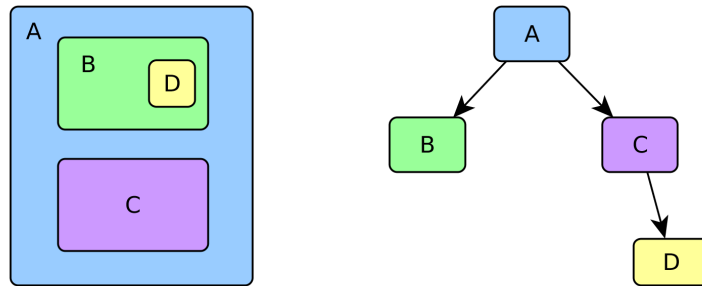
7.1.2 Il database a grafo per IMAGACT

A partire da questa breve classificazione, si vede chiaramente che la maggior parte dei database NoSQL non è adeguata per un'ontologia linguistica come IMAGACT, in cui la quantità di dati è limitata, ma la struttura dati è articolata e deve consentire query complesse. I *key-value store* sono strumenti per la gestione di grandi flussi di informazione in ambienti distribuiti e la strutturazione dei dati è praticamente assente; questo li rende inutilizzabili per IMAGACT. I *column-family store* consentono una rappresentazione dell'informazione più strutturata, ma anch'essi sono più adatti alla gestione di Big Data che di un'ontologia linguistica: in particolare le limitazioni sulla possibilità di incrociare i dati a vari livelli sono un ostacolo allo sfruttamento di IMAGACT. Anche i *Document Database* non sono adeguati perché IMAGACT non ha dati documentali da rappresentare e, inoltre, tutte le informazioni sono molto legate tra loro attraverso relazioni e un limite di questo tipo di modelli è quello di non rendere semplice l'intersezione di documenti. Se quindi queste soluzioni NoSQL non risultano essere appropriate per IMAGACT, i database a grafo hanno invece delle qualità che ne hanno fatto preferire la scelta rispetto ad un RDBMS. Questi infatti, come un database relazionale, consentono di modellare una struttura dati articolata e facilmente interrogabile; oltre a ciò sono in grado di sopperire a quelle criticità che sono emerse nell'implementazione di IMAGACT in un modello relazionale.

Prima di tutto la nuova ontologia basata su un modello dati insiemistico è più facilmente codificabile in un grafo che su tabelle: infatti, come si vede in figura 56, la conversione di un modello insiemistico in un grafo è banale e non necessita di particolari astrazioni. È sufficiente rappresentare gli insiemi come nodi e l'appartenenza ad un insieme con una relazione.

In secondo luogo un database NoSQL risolve il problema di rigidità strutturale dei

Figura 56: Rappresentazione di un modello insiemistico attraverso un grafo



database relazionali che mal si adatta ad un'ontologia linguistica come IMAGACT, che è in continua evoluzione e ha portato ad una esplosione di tabelle. Il database a grafo, non avendo una struttura dati rigida, può essere aggiornato e modificato molto meglio in corso d'opera senza doverne aumentare la complessità strutturale. Un'ulteriore ragione che ci ha fatto preferire il database a grafo rispetto ad un db relazionale è la maggiore vicinanza della struttura dati ad un'ontologia formale RDF. Questo fattore è molto importante per poter produrre una risorsa di rete sfruttabile da algoritmi di elaborazione del linguaggio naturale. Essendo questo uno degli scopi del lavoro di tesi, utilizzare un modello facilmente convertibile in un'ontologia è una qualità desiderabile.

Possiamo dire che in generale una struttura a grafo si adatta bene a formalizzare dati linguistici ed è ampiamente utilizzata in questo ambito. I grafi sono spesso utilizzati per concepire le relazioni tra oggetti linguistici, come per gli alberi sintattici, che modellano le relazioni che intercorrono tra i sintagmi all'interno di una frase o nelle reti semantico-lessicali, che strutturano i significati delle parole in una rete di relazioni semantiche (Nagi, 2013). Inoltre il grafo, più di altre strutture, si adatta ad ospitare l'annotazione linguistica di un testo, che normalmente comprende vari livelli parzialmente o totalmente indipendenti tra loro (Bird and Liberman, 2000).

7.2 IL NUOVO DB

Benché i database a grafo siano una tecnologia recente, essi hanno avuto un forte sviluppo negli ultimi anni e ad oggi è disponibile una grande varietà di questi strumenti⁶. Diversi fattori sono stati presi in considerazione per la scelta del database per IMAGACT. In primo luogo sono stati analizzati gli aspetti relativi all'affidabilità della memorizzazione, cercando soluzioni che potessero offrire le stesse proprietà dei RDBMS, a discapito della scalabilità orizzontale che non è un requisito importante per IMAGACT; fortunatamente molti database a grafo implementano le transazioni ACID. Un altro fattore che ha guidato la scelta è stata la disponibilità di API per diversi linguaggi di programmazione e API REST, in modo da facilitare l'integrazione del database in applicativi software e rendere l'ontologia sfruttabile da algoritmi esterni. In prospettiva di una preservazione del database sul lungo periodo è stato scelto infine di utilizzare solo software open-source e di cui fossero disponibili delle release in produzione. Sono stati trovati 5 database a grafo rispondenti a questi requisiti: ArangoDB⁷, Blazegraph⁸, Neo4J⁹, OrientDB¹⁰ e Titan¹¹. Tra questi abbiamo scelto Neo4J per tre motivi principali: le elevate performance nelle operazioni di ricerca tra nodi (Jouili and Vansteenbergh, 2013), l'adattabilità al dominio delle reti semantiche (Nagi, 2013) e l'espressività del linguaggio di interrogazione (Cypher) che rende particolarmente semplici le operazioni di *retrieval* in IMAGACT.

7.2.1 Neo4J

ARCHITETTURA Neo4J è un database a grafo ad alte prestazioni sviluppato da Neo-Technology, open-source e scritto in Java. Nell'architettura di Neo4J, di cui riportiamo uno schema in Fig. 57, possiamo riconoscere tre livelli, uno basso relativo alla memorizzazione dei dati su disco, uno alto che espone le funzioni per l'interazione con il db ed

⁶ All'epoca delle nostre ricerche (gennaio 2015) Wikipedia (https://en.wikipedia.org/wiki/Graph_database) censiva più di 40 database a grafo disponibili sul mercato.

⁷ <https://www.arangodb.com/>

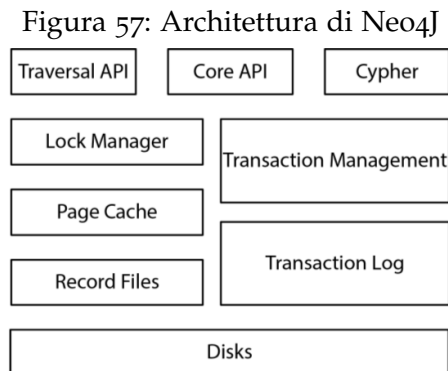
⁸ <https://www.blazegraph.com/>

⁹ <http://neo4j.com/>

¹⁰ <http://orientdb.com/>

¹¹ <http://thinkarelius.github.io/titan/>

uno intermedio che contiene i moduli del Management System.



I dati persistenti sono memorizzati in file indicizzati, gli *store files*, che sono tutti raccolti in un'unica cartella; questo rende particolarmente semplici le operazioni di backup e restore. Un aspetto interessante di Neo4J è l'implementazione nativa della struttura a grafo, cioè gli *store files* ospitano una serializzazione diretta degli elementi del grafo; in molti database invece le strutture di persistenza sono di altro Tipo ed è il DBMS che realizza un'astrazione in forma di grafo. La memoria è gestita in modo efficiente attraverso due cache, una a livello di file system che mantiene delle parti di *store files* sotto forma di record (File System cache), ed una più veloce e di più alto livello che mantiene porzioni di grafo (Object cache). Come abbiamo detto Neo4J supporta transazioni ACID e il Lock Manager si occupa di garantire l'atomicità attraverso un sistema di *lock* per le operazioni di lettura e scrittura. È inoltre possibile preservare lo storico delle transazioni (Transaction Log) fissandone l'estensione temporale nel file di configurazione e questo consente di fare il rollback del database.

L'interazione con Neo4J può avvenire a vari livelli, dipendentemente dalle necessità applicative. Al livello più basso troviamo la Core API, che permette di interagire programmaticamente via Java direttamente con le strutture di base di Neo4J: essa contiene ad esempio la mappatura in oggetti di nodi e relazioni. Sempre via codice è possibile accedere ad una struttura di più alto livello, il *framework* Traversal, che mappa oggetti più astratti (come Path o Branch) ed implementa una serie di algoritmi di attraversamento del grafo. Infine è possibile interagire con il database attraverso query in Cypher, un linguaggio dichiarativo specifico di Neo4J. Sono disponibili numerose API che consentono di effettuare richieste in Cypher da diversi linguaggi di programmazione; è poi possibile

interrogare il database tramite messaggi JSON utilizzando la API REST.

STRUTTURA LOGICA Formalmente un grafo è una coppia $G = (V, E)$ dove V (*vertex*) è l'insieme dei nodi e $E \in V \times V$ (*edge*) è l'insieme degli archi. Sopra questa definizione generale, i grafi possono distinguersi in varie tipologie in dipendenza delle loro caratteristiche (orientati, aciclici, etichettati, ecc.). Il modello dati di Neo4J è il **property graph model**, un grafo orientato in cui sia i nodi che gli archi possono essere arricchiti con delle proprietà. Gli elementi principali del modello sono:

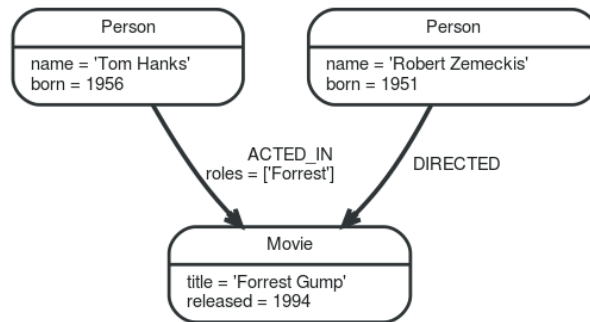
- **nodi**: rappresentano le entità concettuali del modello logico e corrispondono ai nodi del grafo;
- **relazioni**: corrispondono agli archi del grafo, si definiscono tra due nodi e sono direzionate (e quindi non simmetriche); rappresentano le relazioni tra le entità;
- **proprietà**: sono coppie chiave-valore in cui il valore è tipato e possono essere legate sia ai nodi che alle relazioni per specificarne le proprietà interne;
- **etichette**: servono a raggruppare i nodi all'interno di un insieme e determinano quindi la classe del nodo; un nodo può appartenere a una o più classi o anche a nessuna.

Lo schema riportato nella figura 58 è estratto dalla documentazione di Neo4J¹² ed esemplifica chiaramente il ruolo di questi 4 elementi nel modello dati: nell'esempio in questione ci sono 3 nodi, ciascuno dei quali ha un'etichetta (Person e Movie) che ne specifica la classe; i nodi sono collegati tra loro attraverso 2 relazioni, ACTED_IN e DIRECTED; sono presenti delle proprietà (di Tipo numerico o stringa) sia legate ai nodi (name, born, title, released) sia alla relazione ACTED_IN (roles).

Come abbiamo detto precedentemente, un database a grafo ha una struttura dati più flessibile e *schema-less*. Infatti etichette e proprietà sono attributi associabili in modo libero ai nodi senza dover predefinire una struttura; per meglio capire cosa questo significhi in termini operativi, rimaniamo sull'esempio di Fig. 58. Se guardiamo ai due nodi Person, riusciamo immediatamente a capire come è strutturata l'informazione: c'è una classe

¹² <http://neo4j.com/docs/>

Figura 58: Esempio di rappresentazione dell'informazione in Neo4J



generale Person e due istanze (oppure oggetti, oppure individuals) di quella classe con due unità di informazione tipate, che a seconda dell'ambito applicativo che consideriamo potremmo chiamare proprietà, campi, attributi o variabili. Inoltre chi ha un minimo di esperienza con le strutture dati sarà probabilmente portato ad indurre che questa strutturazione corrisponda alla definizione di una Person e che quindi tutte le istanze di Person abbiano due proprietà, *name* e *born* (che magari possono avere valore *NULL*). In realtà nel database a grafo non esistono definizioni di strutture, quindi ad esempio è possibile inserire una Person senza la proprietà *born* (che è molto diverso da *born = NULL*), ma con due proprietà in più, *city* e *job*, che specificano la città di residenza e la professione. Questo può essere fatto senza alterare le due Person già presenti; otterremo quindi istanze della stessa classe con proprietà diverse. Anche il concetto di classe è errato in questo modello, in cui più correttamente questa tipologia di informazione è chiamata **etichetta**. Le etichette infatti determinano soltanto l'appartenenza dei nodi ad un gruppo e possono essere associate in modo libero; quindi è possibile rimuovere un'etichetta da un nodo, oppure aggiungerne di nuove o modificarne il nome senza che questo influenzi il contenuto del nodo. Un discorso analogo può essere fatto per le relazioni e le loro proprietà: è possibile aggiungere un'altra relazione ACTED_IN senza specificare *roles* e viceversa aggiungere una relazione DIRECTED con una proprietà *year* che specifichi ad esempio l'anno in cui è stato girato il film.

Questa assenza di strutturazione rigida dei dati, se da un lato rende estremamente semplice l'utilizzo del database, dall'altro impone un maggior controllo da parte del software per mantenere i dati coerenti. Molti dei vincoli di integrità dei database relazionali che prevengono l'inserimento di dati non conformi al modello strutturale

contrastano con la proprietà *schema-less* che caratterizza i database NoSQL e non sono quindi disponibili in questi strumenti, compresi i database a grafo. Questo discorso è generale e vale anche in Neo4J, in cui però sono supportati due tipi di vincoli, *unique constraints* e *existence constraints*, che permettono di ridurre i controlli di coerenza dei dati a livello di applicazione.

Le *unique property constraints* consentono di definire l'unicità del valore di una proprietà per tutti i nodi che hanno una specifica etichetta. Se ad esempio volessimo identificare univocamente gli elementi della classe Movie attraverso un ID, potremmo dichiarare il campo ID come *unique property* degli elementi che hanno l'etichetta Movie. In tal modo, Neo4J manterrà il campo ID unico, prevenendo la scrittura di un nodo Movie con lo stesso ID di uno già presente. È importante notare che questo vincolo assicura l'unicità di una proprietà, ma non la sua presenza: come abbiamo visto precedentemente, infatti, le proprietà di un nodo non sono un elemento necessario, quindi i nodi per i quali la proprietà non è definita sono sempre accettati. In altri termini una *unique property constraint* garantisce che tutti i valori di una specifica proprietà siano diversi per ogni nodo e non che tutti i nodi abbiano quella proprietà. Rendere una proprietà obbligatoria è comunque possibile utilizzando la seconda tipologia di vincolo, *property existence constraints*, che garantisce la presenza di una proprietà ed è applicabile sia ai nodi che alle relazioni. Sempre in riferimento all'esempio precedente potremmo impostare il vincolo che tutti i Movie debbano avere un titolo; in tal modo il database non accetterà inserimenti di Movie privi di questa proprietà. Questi due vincoli sono gli strumenti che Neo4J mette a disposizione per mantenere la coerenza dei dati; tutti gli altri controlli devono essere implementati a livello di applicazione. Ad esempio non sono presenti vincoli di integrità referenziale ed in generale nessuna restrizione sulle relazioni tra nodi diversi.

7.3 QUESTIONI DI MIGRAZIONE

Come abbiamo detto un'ontologia linguistica in espansione come IMAGACT necessita di una certa flessibilità strutturale, in modo da potersi facilmente adeguare ai cambiamenti del modello e all'aggiunta di nuovi livelli di annotazione. Questo requisito non è derivato

soltanto da ipotesi sulle potenzialità di crescita della base dati, ma anche da evidenti criticità attualmente presenti nel database. In particolare sono emersi due problemi causati dall'utilizzo di un database relazionale e difficilmente risolvibili con l'attuale struttura dati: la rappresentazione dei verbi e delle frasi. Un'altra questione importante è quella della rappresentazione delle scene: l'utilizzo di scene vuote infatti può causare dei problemi di perdita di informazione nel processo di migrazione e deve quindi essere gestito in modo controllato.

7.3.1 *Rappresentazione dei verbi*

Come abbiamo visto nel Par. 3.2, alcune lingue, quelle per cui è stato fatto un lavoro corpus-based, hanno i tipi azionali, mentre le altre, derivanti dall'annotazione CBE, hanno soltanto l'associazione tra verbo e Scena. Per alcuni verbi quindi abbiamo una distinzione tipologica che esprime la variazione verticale del loro significato e per ciascun Tipo una o più scene prototipiche che ne rappresentano l'estensione; per gli altri invece non c'è la suddivisione in tipi e il riferimento azionale è esplicitato in modo diretto, indipendentemente dalla loro variazione. Ci sono poi alcuni verbi che non provengono né da corpus, né dall'annotazione CBE, ma sono stati inseriti in fase di traduzione delle frasi; questi verbi non hanno tipi e non hanno collegamenti diretti con le scene, ma sono soltanto legati a una frase.

In IMAGACT quindi un'unica entità concettuale, il verbo, proviene da tre diverse interfacce di annotazione e si collega a livelli diversi con le altre entità del db; questo ha portato il database ad avere tre rappresentazioni del verbo, corrispondenti ciascuna ad una tabella indipendente: CorpusVerb, CBEVerb e TranslatedVerb. Un tale livello di disgregazione dei dati costituisce sicuramente un problema per lo sfruttamento dell'informazione, specialmente perché l'entità in questione è il verbo, che è una delle unità di base dell'ontologia. Per esempio un'operazione banale come l'elencazione di tutti i verbi di IMAGACT si traduce in una query su tre tabelle con rimozione dei duplicati.

Parallelamente però si riscontra una difficoltà oggettiva nel ripensare un modello relazionale in cui il verbo sia rappresentato in un'unica struttura dati, poiché sia le proprietà delle tre "tipologie" di verbi, sia le loro relazioni con gli altri dati sono molto

diverse: un *CorpusVerb* ad esempio ha dati relativi alla frequenza del verbo nel corpus, le relazioni con le occorrenze da cui è stato estratto e la suddivisione in tipi azionali; un *CBEVerb* non ha nessuna di queste informazioni, ma contiene i collegamenti con le scene e con le loro *Caption*; un *TranslatedVerb* invece ha soltanto il collegamento con la traduzione di una frase standardizzata. Notiamo quindi che queste differenze tra verbi ne rendono quantomeno complessa la rappresentazione in un'unica entità nel modello relazionale.

La rappresentazione del verbo costituisce un caso in cui la flessibilità strutturale di un database a grafo è molto utile, in quanto consente di avere un'unica entità logica (identificata da un'etichetta), in cui però sia le proprietà sia le relazioni con le altre entità possono variare. Questo permette anche di evitare l'inserimento di doppioni: se lo stesso verbo proviene da due annotazioni diverse (ad esempio sia dal corpus che dalla traduzione), basterà aggiungere ad esso le proprietà e le relazioni dell'una e dell'altra.

7.3.2 *Rappresentazione delle frasi*

IMAGACT è anche un database di frasi semplici standardizzate, che spesso provengono da un'occorrenza nel corpus, ma possono anche essere state create dall'annotatore nei casi in cui un certo Tipo di variazione non fosse presente nel corpus; altre frasi ancora provengono dal processo di traduzione che è stato realizzato su una parte del corpus. Ogni standardizzazione è suddivisa in *chunk* testuali, ciascuno dei quali ha un ruolo tematico. Normalmente le standardizzazioni sono legate ad un Tipo azionale: essendo infatti la testimonianza dei campi di variazione di un verbo, è proprio attraverso il loro *clustering* che sono stati concepiti i tipi. Questo legame tra frase e Tipo costituisce un ostacolo alla dinamicità dei tipi azionali: una modifica ai tipi di un verbo renderebbe necessaria una rianalisi di tutte le frasi collegate ai suoi tipi, che dovrebbero infatti essere riassegnate ai tipi corretti. Questo è un problema perché da un lato il numero di frasi, che è estremamente variabile nei diversi tipi, può essere anche molto elevato; dall'altro la suddivisione in tipi azionali è un giudizio non sempre condiviso e suscettibile di revisioni. Pertanto è stato deciso di legare le frasi alle scene, anziché ai tipi, in modo da rendere agile l'eventuale modifica ai tipi azionali. Per fare questa transizione però è

necessaria una rianalisi delle frasi, che al momento non è stata ancora fatta. Utilizzando le proprietà del database a grafo abbiamo deciso di adottare questo approccio in 3 fasi:

1. per la migrazione viene riportato il collegamento tra frasi e tipi e viene aggiunto anche il collegamento tra frasi e verbi (che è automatico, dal momento che il Tipo appartiene a un verbo);
2. quando si operano modifiche ai tipi, viene eliminata la relazione tra Tipo e frasi; rimane soltanto il collegamento tra Tipo e verbo;
3. in fase di rianalisi delle frasi viene effettuato il collegamento tra le frasi e le scene.

Nel caso delle frasi quindi l'assenza di vincoli sulla struttura dei dati ci dà la possibilità di rimuovere e aggiungere relazioni tra nodi in modo dinamico; questo ci consente di concepire un modello dati evolutivo in cui il legame frase-tipo andrà via via a scomparire, mentre invece cresceranno collegamenti frase-scena.

7.3.3 *Le scene vuote*

Il nuovo modello insiemistico non è equivalente al modello originale: nel processo di migrazione infatti avviene una perdita di informazioni. Questo dipende dalla presenza delle scene vuote: se infatti ogni oggetto Scena avesse il riferimento ad almeno un video prototipico, la struttura dati semplificata sarebbe equivalente all'originale e sarebbe possibile passare da un modello all'altro senza perdita di informazioni, utilizzando un algoritmo iterativo. Lo scenario attuale però è ben diverso, perché in IMAGACT la creazione di scene vuote per rappresentare azioni generali è stata una pratica molto utilizzata e attualmente oltre il 20% degli oggetti Scena sono privi del riferimento ad un video.

Nel processo di migrazione ogni Scena vuota nel modello originale può alterare negativamente la struttura dati di uscita in due modi:

1. Eliminazione del Tipo azionale: se la Scena vuota è una Scena "foglia" o comunque non ci sono scene "piene" ad un livello inferiore della gerarchia, allora il Tipo azionale viene eliminato.

2. Appiattimento dell'informazione: se la Scena vuota è "padre" di una sola Scena il Tipo risulta equivalente al Tipo "figlio". Questo comportamento si può verificare anche quando il "padre" ha più di un "figlio", nel caso in cui gli altri figli siano scene vuote e non abbiano alcuna Scena piena nella loro gerarchia ad un qualche livello inferiore.

L'eliminazione del Tipo azionale può essere un problema oppure no, ma naturalmente questo set di casi deve essere analizzato prima di procedere alla migrazione per controllare che ciò che viene eliminato non sia informazione rilevante per l'ontologia. Nell'esempio in Fig. 59 (*attaccare-t4*) la migrazione causa l'eliminazione di tre tipi azionali contenuti in due scene vuote foglia, una che è prototipo di *aderire-t1* e una di *affiggere-t1*. Mentre nel primo caso (*aderire-t1*) l'eliminazione del Tipo è coerente con il contenuto dell'ontologia (*aderire* non è un'azione fisica), gli altri due tipi fanno riferimento ad una Scena propriamente azionale che, per incompletezza, non è rappresentata da un video; in questo caso l'informazione non deve essere eliminata perché rilevante, si deve quindi aggiungere un video prototipico alla Scena vuota prima di procedere alla migrazione.

Anche nel caso di appiattimento dell'informazione deve essere fatto un controllo preliminare per evidenziare i possibili problemi: in alcuni casi l'appiattimento può essere opportuno in favore di una semplificazione, in altri invece può portare a scarsa coerenza nella definizione del Tipo. Nelle figura seguente è mostrato il Tipo 2 di *attaccare* prima e dopo la migrazione. In questo caso l'appiattimento dell'informazione provoca un problema nella definizione del Tipo, che risulta equivalente al Tipo 1 del verbo inglese *sew*. Questo non è corretto perché *attaccare-t2* (che non esprime il modo in cui si attacca una parte a un oggetto) è più generale di *sew-t1* (che descrive l'azione specifica di cucire). Anche in questo caso per superare il problema si deve aggiungere un video prototipico alla Scena vuota di *attaccare-t2*.

Non sempre la presenza di una Scena vuota provoca una perdita di informazione, che avviene solo quando il Tipo non risulta differenziato dall'insieme dei suoi prototipi: nell'esempio precedente di *attaccare-t4* (Fig. 59), la Scena vuota di *stick-T3* non provoca un appiattimento del Tipo, poiché l'insieme delle scene prototipo sotto-ordinate è un differenziale rispetto agli altri tipi. Nel database sono stati trovati 36 tipi azionali (italiani e inglesi) in cui la presenza di scene vuote causerebbe una perdita di informazione in

Figura 59: La strutturazione gerarchica del Tipo *attaccare-T4*

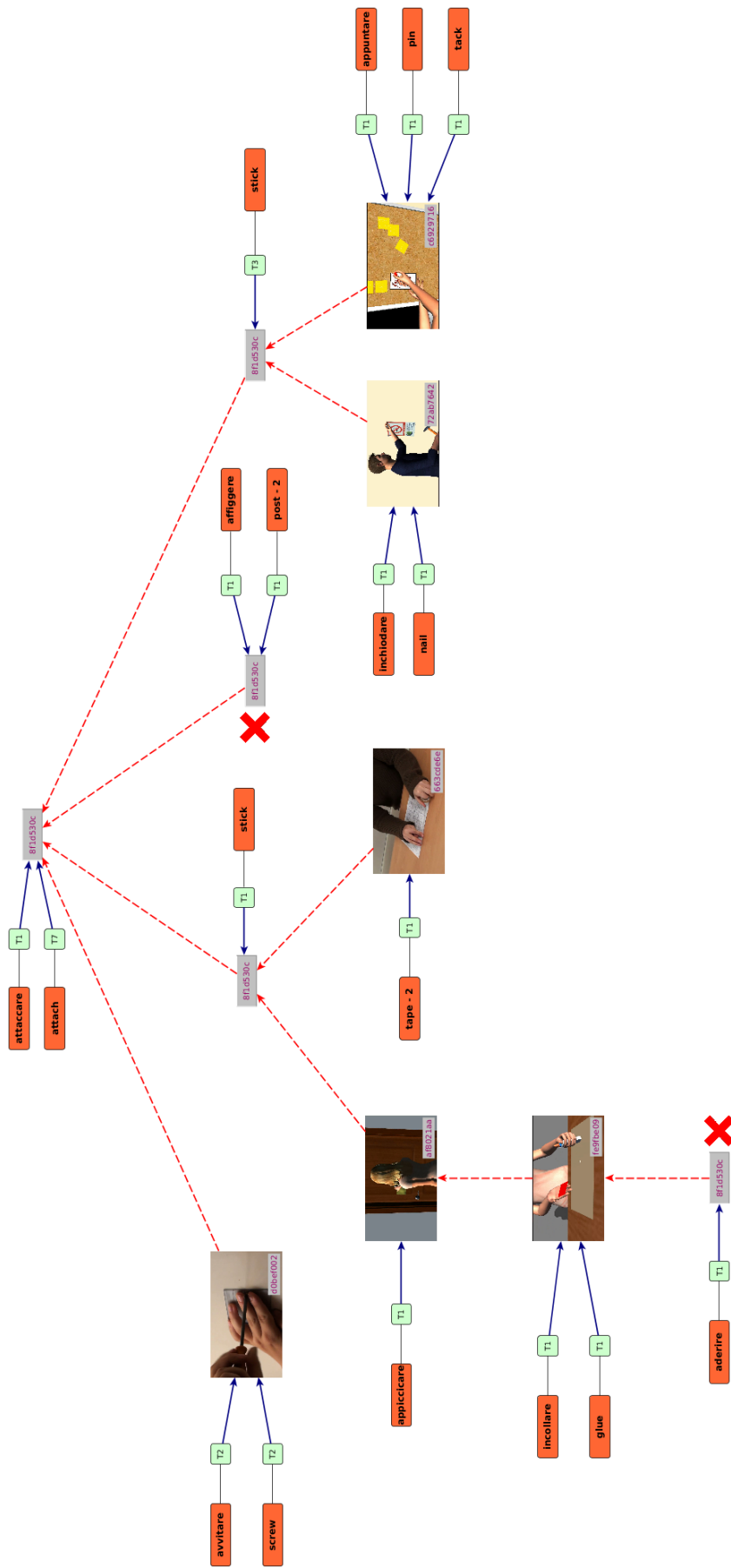
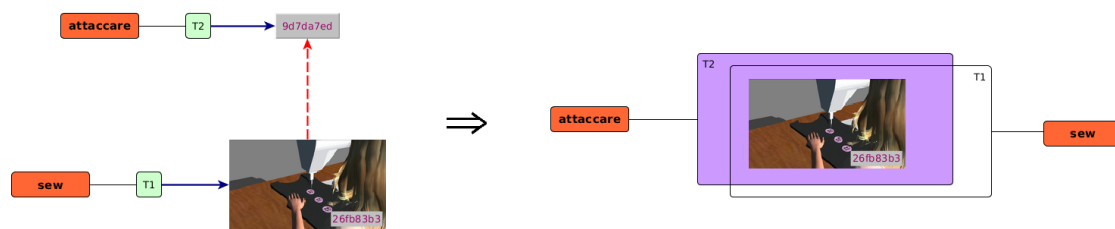


Figura 60: La scomparsa dell'informazione differenziale tra *attaccare* e *sew*

fase di migrazione; questi dati sono stati riaggregati in 4 tipologie di casi critici presenti in IMAGACT ed esemplificati in Figura 61.

I primi due casi portano a un collasso dei tipi, per cui tipi più generali risultano avere la stessa estensione di tipi più specifici. I casi 3 e 4 invece portano ad una scomparsa dei tipi azionali; nel Tipo 4 in particolare viene eliminata l'intera gerarchia.

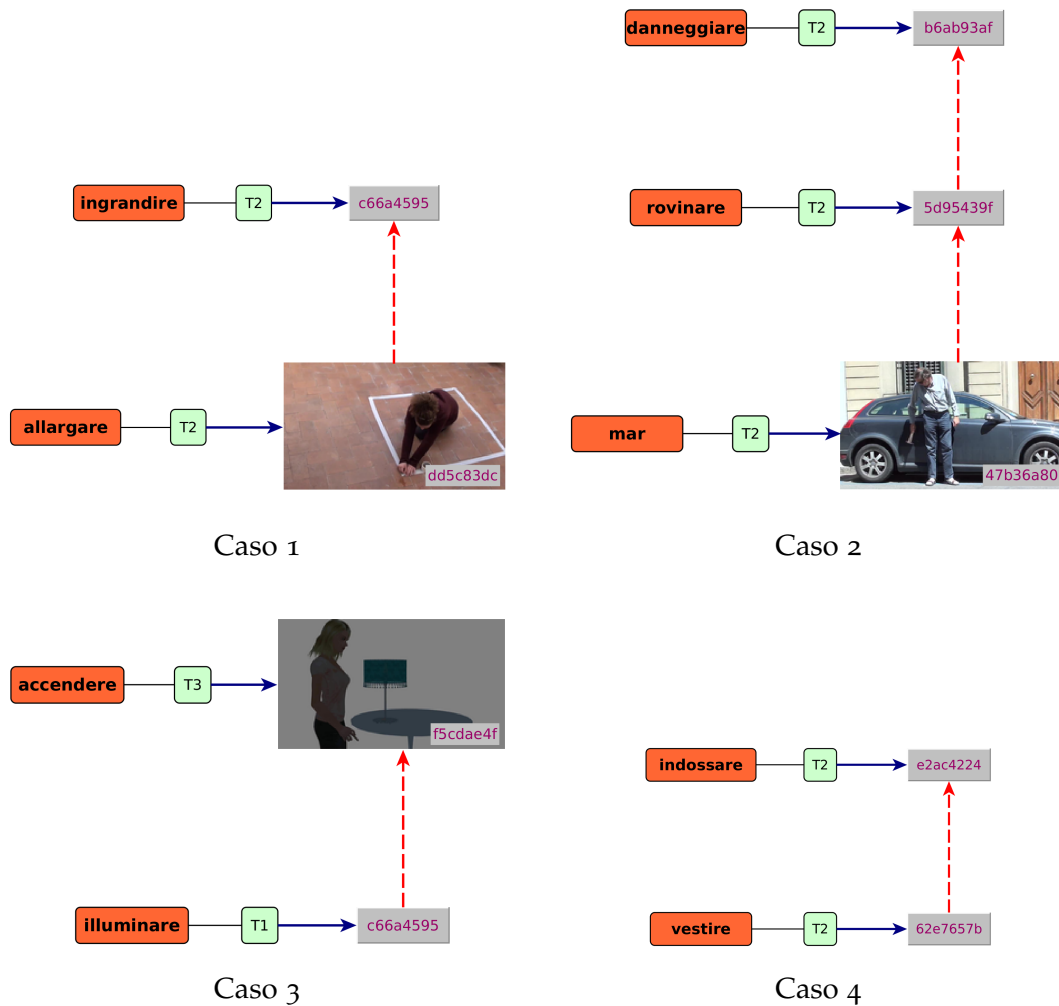
7.4 IMAGACT SU NEO4J

7.4.1 Il database MySQL d'origine

Il modello gerarchico di IMAGACT è implementato su un database relazionale MySQL, utilizzato sia per ospitare l'ontologia, sia per la gestione delle diverse interfacce di annotazione. In questa struttura dati l'informazione linguistica è distribuita in modo sparso sulle 67 tabelle del database e si riscontra una certa lontananza tra il modello dati concettuale e la struttura di persistenza che lo ospita. Questo problema è abbastanza comune nello sviluppo di basi dati linguistiche (Dimitriadis and Musgrave, 2009; Dimitriadis and Musgrave, 2009) e dipende in massima parte dalla costruzione incrementale del modello logico, che deve tenere il passo con l'evoluzione del modello teorico.

Nel caso di IMAGACT la costruzione dell'ontologia ha visto diverse fasi di analisi e annotazione, realizzate in modo per lo più sequenziale e strettamente dipendenti l'una dall'altra: dall'analisi corpus-based delle occorrenze si è passati all'annotazione linguistica su vari livelli, al mapping inter-linguistico dei concetti azionali, alla costruzione di scene prototipiche, fino ad arrivare all'estensione dell'ontologia a nuove lingue. Il modello gerarchico finale derivato da questa complessa procedura di realizzazione è

Figura 61: Casi critici che portano a una perdita di informazione con la migrazione



quello descritto nel capitolo 4 ed è molto diverso dal modello dati concepito in partenza, sia perché più ricco, sia perché modificato in funzione delle necessità derivate dagli avanzamenti del *framework* teorico.

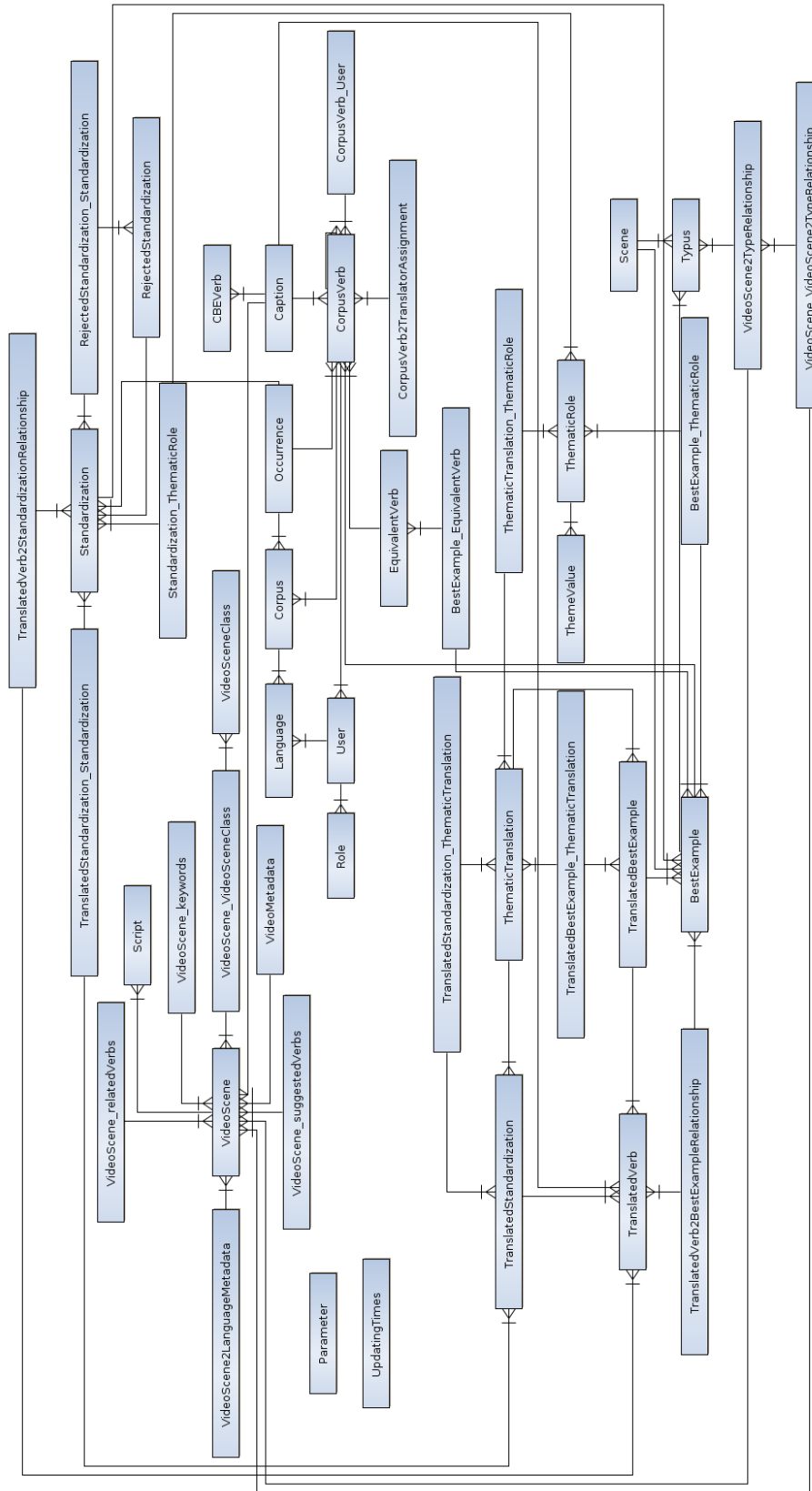
Il cambiamento teorico che ha avuto maggior impatto sulla rappresentazione dei dati è stata la diversa concezione degli oggetti *Scena*, che col tempo hanno acquisito quel ruolo ambiguo a metà tra un prototipo e un concetto azionale (nel database il video e la *Scena* sono entità distinte e fanno riferimento alle tabelle **Scene** e **VideoScene**). Questo si è concretizzato operativamente con l'imposizione dei vincoli (1) e (2) alla relazione di istanza (Par. 4.2), da cui è derivato il modello gerarchico. A causa dell'impossibilità di modifica della base dati in corso d'opera senza una sua riprogettazione, il database

ha mantenuto la struttura originale, in cui l'istanza è, come il prototipo, una relazione tra un Tipo ed una Scena. In questa rappresentazione l'istanza, essendo una relazione tra oggetti diversi, non è transitiva e ciò rende complesse la gestione e lo sfruttamento dell'ontologia. Abbiamo in particolare due problemi:

1. informazione ridondante: ogni istanza è una relazione indipendente, pertanto per poter preservare il vincolo (1) è necessario inserire una nuova relazione tra la Scena e ogni Tipo sopra-ordinato. Questo genera un'esplosione di relazioni difficile da gestire: inserimenti, cancellazioni e modifiche di un'istanza si traducono in una molteplicità di operazioni, una per ogni Tipo sopra-ordinato;
2. difficoltà di formulare inferenze: benché il modello strutturale implementi il modello gerarchico e quindi sia possibile formulare le inferenze, c'è una difficoltà di *retrieval* derivante dalla complessità delle query al database richieste per questo Tipo di operazioni.

Ridondanza di informazione e difficoltà di *retrieval* sono due problemi che non riguardano solo gli oggetti Scena, ma sono diffusi un po' su tutta la base dati. La figura 62 mostra il diagramma relativo alla struttura del database MySQL di IMAGACT, da cui emerge in modo abbastanza chiaro come lo sviluppo incrementale dell'ontologia abbia portato a una rappresentazione dei dati molto disgregata. Possiamo notare ad esempio che l'entità Verbo, che è il fondamento della risorsa linguistica, è memorizzata su quattro diverse tabelle: CorpusVerb, CBEVerb, EquivalentVerb e TranslatedVerb. Ciascuna di esse infatti contiene verbi che hanno diverse informazioni, in quanto provenienti da fasi di lavoro separate: rispettivamente l'estrazione di dato da corpus, l'estensione ad altre lingue, l'annotazione dei tipi azionali e la traduzione delle frasi. Questo ha inoltre introdotto un problema di duplicazione dei dati, in quanto uno stesso lemma verbale può ad esempio essere stato inserito sia come verbo di traduzione, sia come verbo equivalente in fase di annotazione.

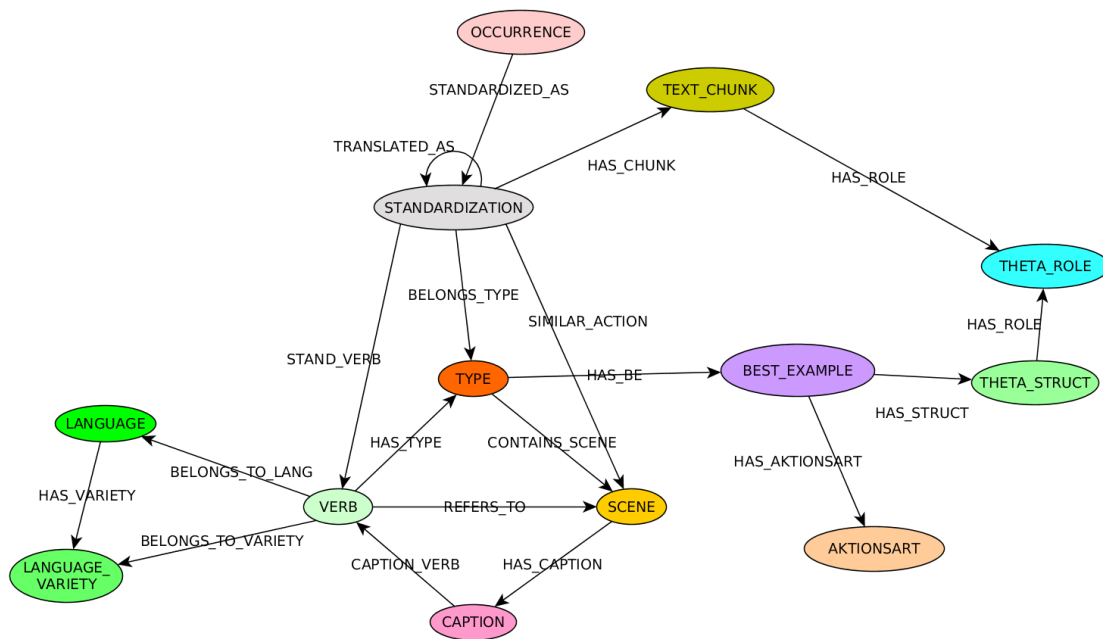
Figura 62: Schema logico del database MySQL di IMAGACT



7.4.2 Implementazione della struttura dati in Neo4J

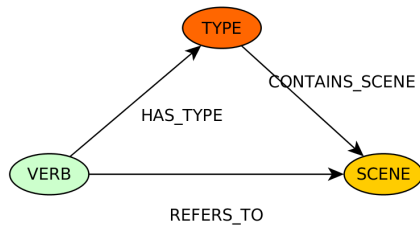
Il modello dati insiemistico progettato per IMAGACT è stato tradotto in una struttura dati a grafo in grado di rappresentare in maniera compatta i diversi livelli di annotazione dell'ontologia (Fig. 63).

Figura 63: La struttura dati implementata su Neo4J



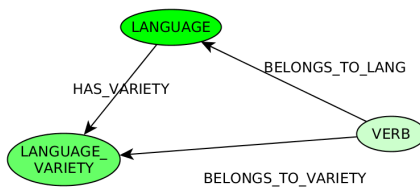
Di seguito riportiamo la descrizione dettagliata della struttura dati e le scelte fatte per modellare le varie entità di IMAGACT, iniziando dal cuore dell'ontologia, ossia i rapporti tra verbi, tipi azionali e scene prototipiche.

Per i verbi di lingue per cui è stata fatta un'annotazione corpus-based abbiamo una distinzione tipologica che esprime la variazione verticale del loro significato e per ciascun Tipo una o più scene prototipiche che ne rappresentano l'estensione; per gli altri invece il riferimento azionale è esplicitato in modo diretto, senza una classificazione dalla loro variazione.



Un verbo si riferisce a una o più scene
 Un verbo può avere 0, 1 o più tipi
 Un Tipo si estende a una o più scene

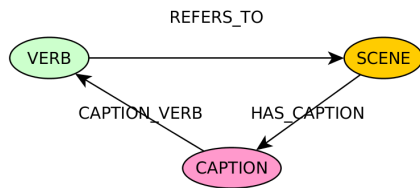
È impostata una relazione REFERS_TO tra i Verb e ogni Scene a cui questi si possono riferire. Per i verbi che hanno tipi vengono creati tanti nodi Type quanti sono i tipi del verbo; il Verb è collegato ai suoi Type tramite HAS_TYPE ed è impostata una relazione CONTAINS_SCENE tra i Type e le azioni a cui questi si estendono. Ogni verbo appartiene ad una lingua o eventualmente ad una varietà linguistica: come abbiamo visto infatti la variazione del lessico verbale in alcune lingue (come lo spagnolo o il portoghese) è molto forte in dipendenza della varietà linguistica.



Un verbo appartiene a una e una sola lingua
 Una lingua può avere delle varietà linguistiche
 Un verbo può appartenere a una varietà linguistica

Ogni Verb è collegato a una e una sola Language tramite BELONGS_TO_LANG; c'è una relazione HAS_VARIETY tra una Language e ogni sua LanguageVariety; se il verbo è specifico di una varietà si imposta una relazione BELONGS_TO_VARIETY tra il Verb e la LanguageVariety.

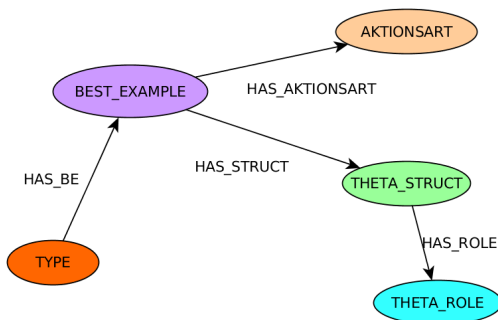
Ogni Scena è descritta da un certo numero di frasi semplici (*caption*), una per ogni verbo che può riferirsi alla Scena.



Una Scena è descritta da una o più Caption
Una caption utilizza un verbo

Per ogni Scene è inserita una relazione HAS_CAPTION con tutte le Caption che la descrivono; per ogni Caption c'è una relazione CAPTION_VERB con il Verb della Caption.

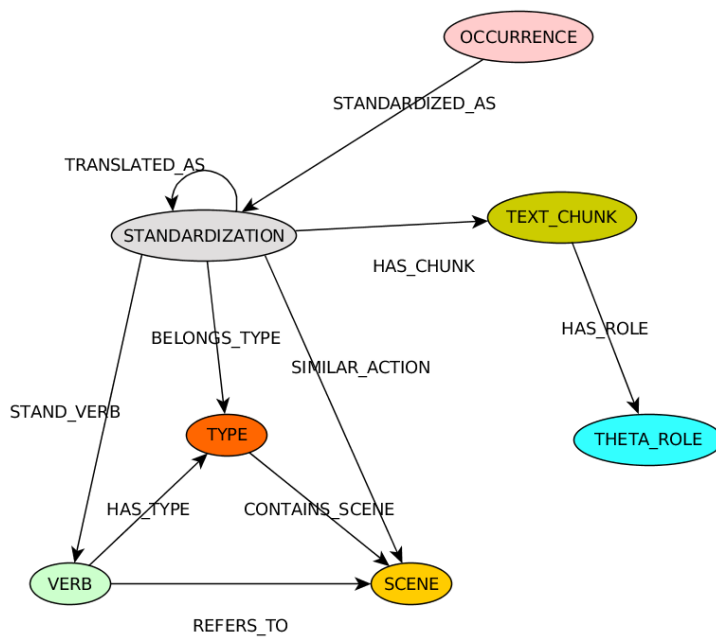
Ogni Tipo azionale ha uno o più BestExample, che ne specifica le variazioni di struttura tematica e aktionsart. Un BestExample ha sempre una aktionsart (processo, evento o stato) ed una struttura tematica, che corrisponde ad una sequenza di ruoli tematici.



Un Tipo ha uno o più BestExample
Un BestExample ha una aktionsart
Un BestExample ha una struttura tematica
Una struttura tematica è formata da ruoli tematici

C'è una relazione HAS_BE tra un Type e tutti i suoi BestExample; ogni BestExample ha un'unica relazione HAS_AKTIONSART con una AKTIONSART; ci sono 3 nodi di Tipo Aktionsart: Process, Event or Protracted Event e State. C'è una relazione HAS_STRUCT tra un BestExample e una ThetaStruct; una ThetaStruct ha una relazione HAS_ROLE con ogni ThetaRole da cui è formata; HAS_ROLE ha una proprietà position con valore numerico incrementale che specifica la posizione del ThetaRole nella ThetaStruct; Ci sono 18 nodi ThetaRole, uno per ogni ruolo tematico definito in IMAGACT.

Le standardizzazioni possono essere state create a partire da occorrenze del corpus (la maggior parte) o direttamente dall'annotatore sulla base della sua competenza; per alcune standardizzazioni è presente una traduzione. In generale tutte le standardizzazioni sono annotate con i ruoli tematici. Inoltre una standardizzazione è sempre legata ad un verbo e può opzionalmente fare riferimento ad un Tipo (nella fase iniziale di migrazione) o a una Scena (dopo una revisione).



Un'occorrenza può avere una standardizzazione	Una standardizz. è legata ad un verbo
Una standardizzazione può avere una traduzione	Una standardizz. può essere legata a un Tipo
Una standardizzazione è divisa in chunk	Una standardizz. può essere legata a una Scena
Un chunk ha un ruolo tematico	

Se l'occorrenza del corpus è standardizzata è impostata una relazione STANDARDIZED_AS tra una Occurrence e una Standardization; se la frase è tradotta è impostata una relazione TRANSLATED_AS tra le due Standardization; c'è una relazione HAS_CHUNK tra una Standardization ogni suo TextChunk; HAS_CHUNK ha una proprietà position, con valore numerico sequenziale, che specifica la posizione del TextChunk all'interno della Standardization. Ogni TextChunk ha una relazione HAS_ROLE con il ThetaRole corrispondente. Ogni Standardization ha una relazione STAND_VERB con un Verb; Una Standardization può essere legata a un Type con la relazione BELONGS_TYPE; una Standardization può essere legata a una Scena con la relazione SimilarAction.

7.4.3 Migrazione del DB

La migrazione di IMAGACT è stata condotta in modo semi-automatico; sono stati realizzati (in ordine):

- script SQL con le query per estrarre i dati dal database MySQL in formato CSV;
- revisione manuale dei file CSV per uniformare i dati, correggere gli errori e eliminare le informazioni ridondanti;
- script Cypher con le query per caricare i dati in Neo4J.

La fase della revisione manuale è stata necessaria per risolvere i problemi descritti all'inizio di questo capitolo e derivati dall'evoluzione della struttura dati sopra un database relazionale. Non è stata quindi realizzata un'unica procedura di migrazione automatica per non introdurre rumore nel nuovo database. Sono stati migrati i dati linguistici dell'ontologia, mentre non sono state considerate le tabelle di Tipo tecnico, cioè quelle utilizzate dalle varie interfacce di annotazione di IMAGACT per memorizzare informazioni di sessione, opzioni di visualizzazione, ecc.

Il database MySQL di IMAGACT consta di 67 tabelle così suddivise (Tab. 3):

- 43 tabelle contenenti dato linguistico;
- 9 tabelle che gestiscono le funzionalità delle interfacce di annotazione;
- 5 tabelle di appoggio contenenti dati temporanei per accelerare le funzioni di ricerca;
- 10 tabelle relative a funzionalità non ancora implementate.

A loro volta nelle 43 tabelle di contenuto linguistico ci sono:

- 3 tabelle vuote, per le quali era stato previsto un ruolo, ma non sono mai state utilizzate;
- 1 tabella dismessa, che era stata utilizzata in una fase iniziale, ma poi abbandonata;

- 6 tabelle non informative, che contengono cioè dati ridondanti (presenti in altre tabelle) oppure non significativi (ad esempio commenti e note dell'annotatore relativi al lavoro *in progress*).

Le restanti 33 tabelle contengono effettivamente l'informazione ontologica rilevante che è stata presa in considerazione per la migrazione. Per adesso 21 tabelle sono state migrate all'interno di Neo4J: abbiamo infatti temporaneamente escluso due dati, quelli relativi ai verbi equivalenti (3 tabelle) e quelli relativi alle traduzioni (9 tabelle); per entrambi le motivazioni riguardano la pulizia del dato immesso nel nuovo database. I verbi equivalenti nel db di IMAGACT sono dei verbi inseriti dall'annotatore nella fase di creazione dei tipi che (a) all'interno del Tipo sono tra loro localmente equivalenti e (b) nel loro insieme costituiscono un differenziale linguistico rispetto agli altri tipi del verbo annotato. Questi verbi sono stati uno dei dati in ingresso sfruttati per la fase di mapping, nella quale si sono costruite le relazioni tra verbi, tipi e scene che hanno dato origine all'ontologia; i verbi equivalenti appartengono quindi ad una fase iniziale dell'annotazione. D'altra parte però corrispondono a giudizi di competenza sull'estensione dei tipi azionali e potrebbero quindi risultare utili in una revisione per colmare delle mancanze dell'ontologia. Sebbene avessimo potuto aggiungere l'informazione al database, inserendo una relazione tra verbo e Tipo "più debole" di HAS_TYPE, abbiamo deciso di non migrare questo dato, sulla base di alcune considerazioni:

- molto spesso i verbi equivalenti non contengono lemmi verbali, ma perifrasi, inserite per creare il differenziale linguistico (in tutti quei casi in cui non si sono trovati verbi localmente equivalenti);
- i verbi equivalenti sono stati inseriti da più annotatori in campi di testo libero e contengono una certa quantità di rumore (doppi spazi, lettere maiuscole, errori di battitura);
- benché l'interfaccia contenesse una checkbox per discriminare le perifrasi dai lemmi verbali, molti di questi valori sono sbagliati (lemmi verbali salvati come perifrasi e perifrasi salvate come lemmi verbali).

Alla luce di questi fatti abbiamo ritenuto opportuno di non migrare i verbi equivalenti, che potranno essere inseriti nel nuovo database soltanto dopo una revisione e una pulizia dei dati.

Le traduzioni presenti in IMAGACT sono un piccolo corpus di frasi parallele italiano-inglese che comprende circa ventimila frasi semplici. Questa risorsa è derivata dall'ontologia, dal momento che le frasi tradotte sono una parte delle standardizzazioni di IMAGACT, ma è indipendente da essa. In particolare le possibilità di sfruttamento sono molto diverse ed è allo studio la realizzazione di una memoria di traduzione (TM), in modo da poter fungere da training set per algoritmi di Machine Translation o da ausilio in applicazioni di traduzione assistita. Inoltre anche le traduzioni hanno un problema di pulizia dei dati, in questo caso causato da alcuni bug che erano presenti inizialmente nell'interfaccia di traduzione: questo è portato a una scrittura errata dei record nel database. Attualmente è in corso una revisione puntuale del corpus di frasi parallele, col fine di rimuovere il rumore e produrre una risorsa pulita e sfruttabile. Per questi motivi le traduzioni sono state escluse dalla migrazione, almeno in questa prima fase.

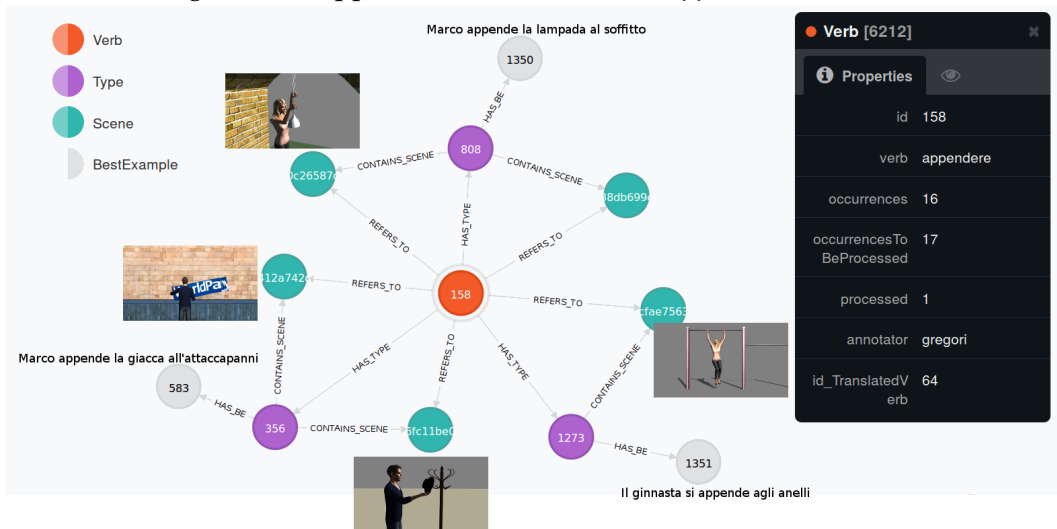
L'estrazione dei dati dal database MySQL è avvenuta attraverso l'esecuzione di 21 query SQL incapsulate in uno script bash, che attraverso il client *mysql* interroga il database ed esporta tutti i dati in file CSV. Per il caricamento dei dati sul database Neo4J abbiamo implementato un tool Java, in grado di effettuare il parsing dei file CSV, la compilazione delle query Cypher di inserimento e la loro esecuzione sul server di Neo4J. Per quest'ultimo task ci siamo avvalsi della API Java messa a disposizione dal database.

Tra le due procedure di estrazione e inserimento è stato necessario inserire una fase manuale di controllo sui dati, per rimuovere alcune incoerenze e le informazioni duplicate. I verbi, ad esempio, a causa della loro distribuzione sparsa su più tabelle nel database, contenevano molti duplicati e, poiché sono tutti dati testuali provenienti da annotazioni, spesso erano duplicati non esatti e quindi difficili da trovare automaticamente. Alcuni esempi sono verbi inglesi riportati una volta con il "to" davanti e una volta senza ("hang" e "to hang"), differenze di maiuscole e minuscole ("mangiare" e "Mangiare") o errori di battitura ("attach" e "atach"). Altri esempi di problemi tecnici che inficiano la qualità dei dati sono la presenza di record incompleti, causati da una annotazione non finita, e la gestione non uniforme dei campi di testo vuoti (a volte rappresentati come stringa

vuota, altre volte come NULL). La revisione dei dati volta a rimuovere questi problemi è stata condotta analizzando i file CSV; la scelta di questo formato è stata guidata dalla compatibilità con i fogli di calcolo, in cui le funzionalità di ricerca e ordinamento rendono efficiente l'analisi dei dati.

A livello esemplificativo riportiamo la strutturazione di un verbo generale nel nuovo modello dati (Fig. 64).

Figura 64: Rappresentazione del verbo *appendere* in Neo4J



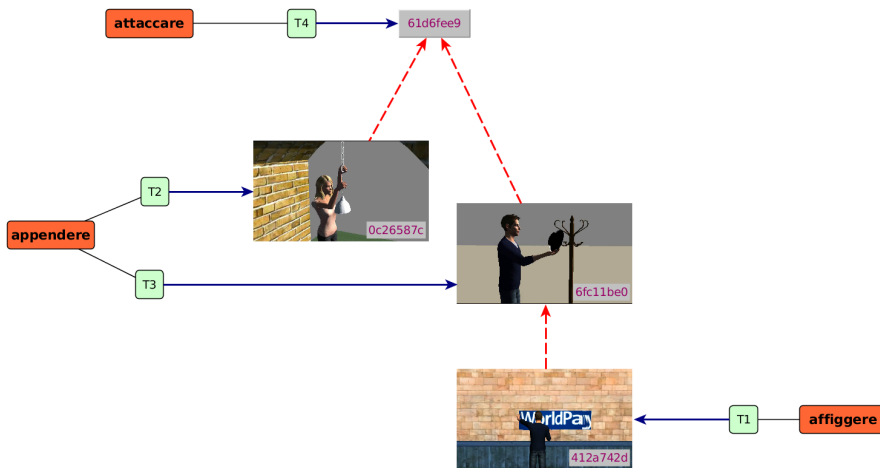
Il verbo italiano *appendere*, con ID 158, ha tre tipi azionali (ID 808, 356 e 1273) che si collegano a 5 scene. La Scena con ID "38db699c" è una Scena vuota, quindi non contiene video; queste scene sono oggetti della vecchia struttura dati e dovranno essere eliminate, ma in questa fase iniziale sono state migrate per non rischiare di perdere informazione rilevante. Come si vede dalla tabella delle proprietà (a destra), il verbo *appendere* proviene da corpus (è un `CorpusVerb`) ed ha proprietà relative al numero di occorrenze nel corpus e all'annotazione. *appendere* era stato inserito anche come verbo di traduzione (nella tabella `TranslatedVerb`), per cui tra le proprietà è riportato l'ID di riferimento al `TranslatedVerb`, in modo da poterci facilmente collegare le frasi tradotte, qualora si decidesse di farlo. Il Tipo 356 contiene due scene prototipo, che in questo modello logico non hanno ruoli diversi; nel modello gerarchico di provenienza invece, solo la Scena "6fc11be0" è il prototipo del Tipo, mentre la Scena "412a742" è il prototipo

di un Tipo del verbo *affiggere*, che è sotto-ordinato rispetto a questo Tipo di appendere. Allo stesso modo sia il Tipo 356 che il Tipo 808 di appendere erano entrambi “figli”, nel modello gerarchico, di un Tipo più generale di *attaccare* (Tipo 529), che in questa struttura ha quindi 3 scene prototipiche: quelle contenute in *appendere-356* e quelle contenute in *appendere-808* (Fig. 65). La Fig. 66 chiarisce i rapporti gerarchici che sussistevano tra i tipi nella vecchia struttura dati.

Figura 65: Rappresentazione del verbo *attaccare* in Neo4J



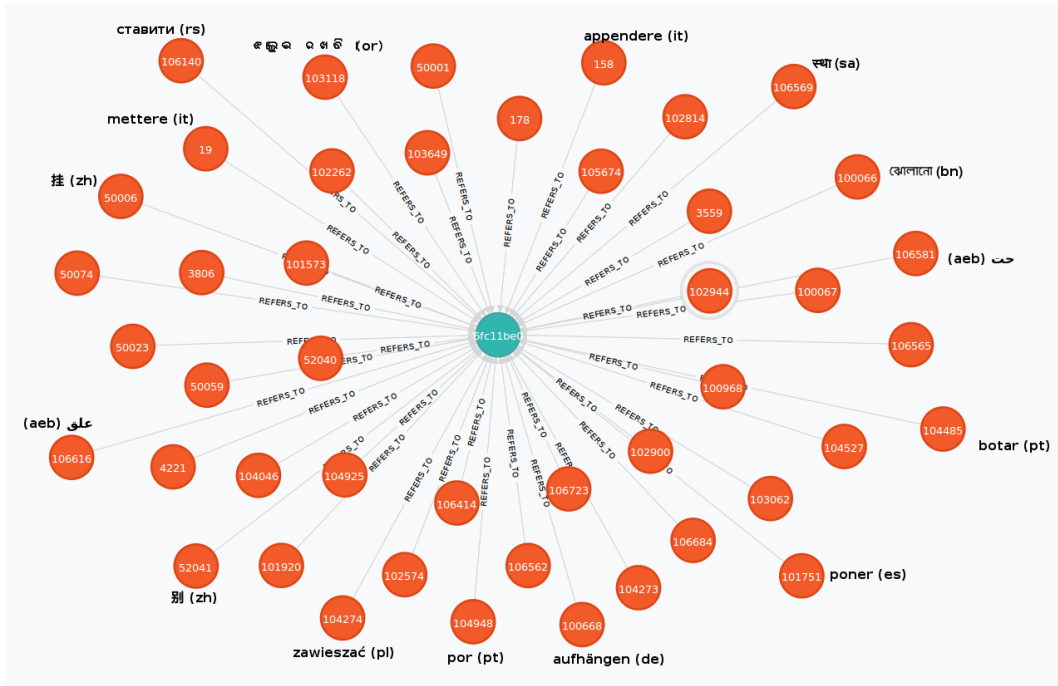
Figura 66: Relazioni tra *attaccare* e *appendere* nel modello gerarchico



In generale una Scena può essere prototipo di tipi diversi nella stessa lingua e in lingue diverse; inoltre può essere riferita direttamente da verbi di lingue che non hanno la suddivisione in tipi azionali. La figura seguente mostra come la Scena “61c11be0”,

identificata dalla Caption “Fabio appende il cappello all’attaccapanni”, sia il riferimento di 45 verbi nelle varie lingue implementate in IMAGACT.

Figura 67: I verbi collegati a una Scena di *appendere* nelle diverse lingue in Neo4J



Ogni BestExample contiene l’informazione relativa a struttura tematica e aktionsart. Come si vede in Fig. 68 il BestExample 1351 (Il ginnasta si appende agli anelli) ha una relazione con una delle 3 aktionsart disponibili (Event or protracted event) ed una con la struttura tematica ACTOR-VERB-DESTINATION, che a sua volta è formata da 3 ruoli tematici ordinati (la proprietà *pos* della relazione HAS_ROLE mantiene l’ordinamento). Tutte le standardizzazioni che fanno parte di questo BestExample vi si collegano con la relazione BELONGS_BE (in questo caso soltanto 2).

Figura 68: BestExample e Aktionsart in Neo4J

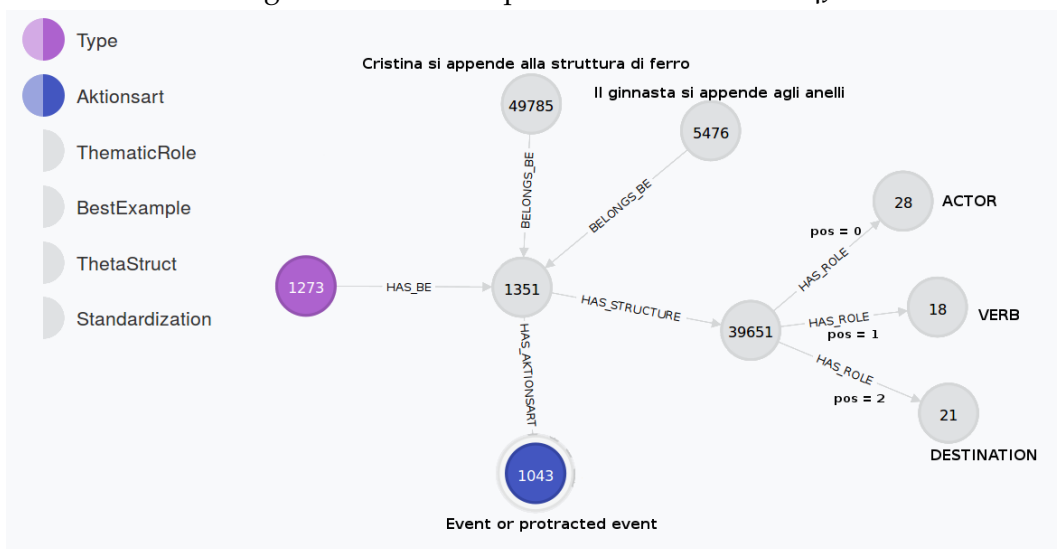


Tabella 3: Le tabelle del database di IMAGACT

Dato linguistico	Nuove funzionalità
BestExample	AbstractVerb
BestExample_ThematicRole	CognitiveMetaphorTag
Caption	CognitiveMetaphorTag_subtags
Caption2LinguisticVariance	GeneralVerb
CBEVerb	GeneralVerbEquivalence
CorpusVerb	MarkedTypus
Language	MarkedTypus_CorpusVerb
LinguisticVariance	MarkedTypus_MarkedVerb
Occurrence	MarkedTypus_VideoScene
Standardization	
Standardization_ThematicRole	
ThematicRole	
ThemeValue	
TranslatedVerb	
TranslatedVerb2BestExampleRelationship	
VideoScene	
VideoScene2TypeRelationship	
VideoScene_VideoScene2TypeRelationship	
BestExample_EquivalentVerb	
Corpus	
CorpusVerb_classes	
EquivalentVerb	
EquivalentVerb_BestExample	
RejectedStandardization	
RejectedStandardization_Standardization	
Scene	
ThematicTranslation	
ThematicTranslation2ThematicRoleRelationship	
ThematicTranslation_ThematicRole	
TranslatedBestExample	
TranslatedBestExample_ThematicTranslation	
TranslatedStandardization	
TranslatedStandardization_Standardization	
TranslatedStandardization_ThematicTranslation	
TranslatedVerb2StandardizationRelationship	
Typus	
VideoMetadata	
VideoScene2LanguageMetadata	
VideoSceneClass	
VideoScene_keywords	
VideoScene_relatedVerbs	
VideoScene_suggestedVerbs	
VideoScene_VideoSceneClass	
Interfacce di annotazione	Tabelle di sistema
CorpusVerb2TranslatorAssignment	VideoScene2MarkedTypusRelationship
CorpusVerb_User	DeNormalizedGlobalCbeReporting
Parameter	DeNormalizedVerb2TypeAssignment
Preference	DeNormalizedVerb2TypeAssignmentComparison
Role	TempEntity
Script	UpdatingTimes
User	
User_CorpusVerb	
User_VideoScene	

VALIDAZIONE DELL'ONTOLOGIA

8.1 IL QUADRO DELLA VALIDAZIONE DI IMAGACT

Una validazione vera e propria di IMAGACT su un task di Word Sense Disambiguation (WSD) prevede un training e un test set. Le occorrenze provenienti dal corpus che popolano il DB di IMAGACT, però, sono troppo sparse per poter addestrare un algoritmo: esse infatti si addensano in pochi tipi azionali, che risultano molto ben rappresentati dal dato testuale, mentre altri tipi contengono poche o pochissime occorrenze; questo fatto è sistematico su tutta l'ontologia e anche su alcuni tipi dei verbi molto frequenti (*mettere, prendere, take, ecc.*), che hanno tipi con solo due o tre testimonianze testuali. Tale insufficienza può essere dimostrata chiaramente. La tabella seguente (Tab. 4) mostra la distribuzione delle occorrenze nei tipi dei primi 5 verbi dell'italiano e dell'inglese in ordine di frequenza. Sono state selezionate soltanto le standardizzazioni delle occorrenze, in modo da escludere dal conteggio gli usi marcati e le occorrenze indecidibili del corpus, per le quali l'annotatore non è stato in grado di fornire un'interpretazione. Sulla base del numero di standardizzazioni presenti in ogni Tipo sono state considerate 5 classi:

- **null** (<5): numero di Tipi con meno di 5 standardizzazioni; la rappresentazione della variazione orizzontale è quasi inesistente;
- **scarso** (5-10): numero di Tipi con 5-10 standardizzazioni; rappresentazione povera della variazione orizzontale;
- **medio** (10-20): numero di Tipi con 10-20 standardizzazioni, che abbiamo considerato un livello medio di copertura della variazione orizzontale;
- **buono** (20-50): numero di Tipi con 20-50 standardizzazioni, considerato un buon campione per rappresentare la variazione orizzontale;

- **ottimo** (50+): numero di Tipi con oltre 50 standardizzazioni: ottima rappresentazione della variazione orizzontale.

Tabella 4: Rappresentazione dei Tipi azionali nel corpus

Verbo	Tipi	Occorrenze	Standard.	<i>null</i>	<i>scarso</i>	<i>medio</i>	<i>buono</i>	<i>ottimo</i>
mettere	16	4018	933	1	2	4	5	4
prendere	16	2459	481	4	6	2	3	1
portare	8	1509	482	1	1	1	1	4
passare	6	946	140	1	1	2	2	0
aprire	9	750	258	3	2	1	1	2
take	13	4006	797	0	1	4	1	7
put	13	3231	928	1	2	1	5	4
move	7	1004	271	2	0	1	2	2
turn	10	702	153	4	1	2	3	0
play	7	615	161	0	2	1	4	0

Da questi dati vediamo che il numero di Tipi con una rappresentazione insufficiente in termini di frasi (*scarso* + *null*) è consistente: 20 Tipi su 55 per l'italiano (~36%) e 14 Tipi su 50 per l'inglese (28%). In sintesi per un Tipo su 3 non ci sono sufficienti dati da corpus perché questi possano essere sfruttati come dataset di addestramento e valutazione di algoritmi di disambiguazione.

La variabilità di popolamento dei tipi appena illustrata dipende dagli usi effettivi dei verbi nella lingua parlata, e dalla probabilità con cui certi tipi di azioni sono riferiti nell'uso linguistico quotidiano ed ha quindi un grande valore rappresentativo dell'uso linguistico, ma purtroppo rende il campione inadatto a fungere da training set. Ipotesi di ricampionamento non sono percorribili, poiché molti tipi occorrono raramente, non sono quindi rappresentati a sufficienza nell'uso linguistico e una loro sufficiente documentazione richiederebbe corpora di dimensioni non raggiungibili allo stato delle tecnologie di *corpus building*.

La validazione dell'ontologia e la verifica del suo potenziale per fini di disambiguazione e traduzione automatica, allo stato, può quindi essere solo indiretta. In questo capitolo riporteremo i risultati di più validazioni condotte sia durante la costituzione dell'ontologia sia a valle della sua realizzazione, specificamente in questo progetto di tesi. Le validazioni in questione sono rivolte in primo luogo a garantire che la metodologia per la selezione

dei dati rilevanti all'induzione da corpus dell'ontologia azionale è fondata. In 8.2.1 vedremo che è validata in particolare la selezione negativa dei dati del corpus tutte le volte l'uso dei verbi non è pertinente all'induzione dei concetti azionali (usi metaforici e astratti). In particolare verrà mostrato che esiste un largo accordo sulla decisione se l'occorrenza di un verbo rappresenta, nel significato proprio del verbo, una azione fisica, o se al contrario l'uso linguistico in quella occorrenza è da considerarsi metaforico. Questa validazione è di particolare importanza perché, oltre a garantire la validità del data set su cui si basa IMAGACT, evidenzia la possibilità di separare nettamente i significati propri dai significati traslati, possibilità che non è implementata nelle ontologie linguistiche correnti quali BabelNet e WordNet. Come vedremo, in particolare in 8.3, da questa mancata distinzione consegue maggiore vaghezza nella definizione dei riferimenti dei verbi, scarso valore interlinguistico dei concetti stessi e ultimamente minori possibilità di disambiguazione. In altri termini i significati astratti sono sia di difficile definizione che scarsamente produttivi cross-linguisticamente (Moneglia et al., 2014a). In 8.2.2 riporteremo i dati che validano la rappresentazione per immagini dei concetti azionali rispetto a task di traduzione assistita. Tale validazione assicura sul fatto che il potenziale di disambiguazione offerto da IMAGACT è in effetti significativo per gli utenti umani, che sono in grado di scegliere in modo univoco e sicuro a quale azione si riferisce una frase semplice, fra le azioni che un verbo può designare, specificamente quando queste sono rappresentate attraverso scene prototipiche. Questo dato è rilevante anche perché segnala uno scarto tra la possibilità di discriminazione univoca dei significati raggiungibile sulla base dei prototipi di IMAGACT e i risultati largamente insoddisfacenti ottenibili con la disambiguazione rispetto ai sensi di WordNet, che come si diceva, risultano sottodeterminati anche per i soggetti umani (Brown et al., 2010; Ng et al., 1999). Entrambe queste validazioni, realizzate da Gagliardi (2014), fanno uso delle metriche standard in questo dominio e in particolare della metrica di *inter-rater agreement* (k di Cohen e k di Fleiss; vedi oltre) standard de facto in questo dominio a partire a Carletta (1996) All'interno del progetto IMAGACT è stato poi misurato il potenziale dei concetti azionali rappresentati nell'ontologia per finalità di riferimento interlinguistico all'azione, ovvero è stata valutata la significatività dei prototipi di azione come termini validi per disambiguare simultaneamente più lingue, ovvero la loro *language independence*.

Tale validazione, detta “verifica di produttività”, dà una misura dell’affidabilità delle previsioni di traduzione di verbi di lingue diverse rispetto a ciascuna azione individuata nell’ontologia ed è stata effettuata in dettaglio per quanto riguarda l’italiano, l’inglese e il cinese. La validazione è avvenuta attraverso un procedimento di verifica fattuale su tutte le occorrenze dei verbi nei corpora che hanno costituito la base per l’induzione di ciascun concetto. Un concetto si ritiene valido per il riferimento interlinguistico se il concetto assicura una traduzione produttiva, ovvero se la verificata sulla base dei dati disponibili è aderente alle seguenti condizioni:

- tutte le frasi di un “language corpus A” che istanziano il concetto (un Tipo di IMAGACT) possono essere identificate da almeno uno stesso verbo in una “lingua target B” ;

e, di converso:

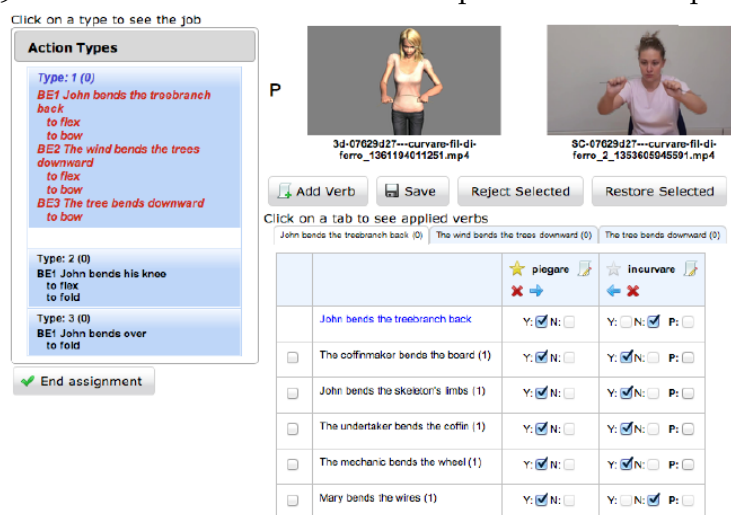
- se tutte le frasi che istanziano il concetto in un “language corpus B” possono essere identificate da almeno uno stesso verbo in una “lingua target A”.

Se entrambe le condizioni sono soddisfatte si prevede conseguentemente che i due verbi possono essere liberamente tradotti in tutte le istanze del concetto, che risulta quindi valido per il riferimento interlinguistico.

La procedura prevede che ogni verbo inseriti in IMAGACT debba essere confrontato con ogni occorrenza del verbo stesso in tutti i tipi di azione a cui si applica. Nel caso illustrato nella figura seguente, l’annotatore, parlante madrelingua italiano, deve scegliere quale verbo traduce *bend* nel Tipo 1, il cui primo esempio è “John bends the treebranch”. L’interfaccia suggerisce il lemma italiano mappato nell’ontologia sullo stesso Tipo (*piegare*) e l’annotatore verifica se tale verbo è applicabile all’intero set di frasi derivate dall’analisi dei corpora e raggruppate come istanze di quel Tipo. L’interfaccia mostra queste frasi listate in sequenza e richiede un giudizio “si/no” sull’applicabilità del verbo previsto. In questo caso il Tipo risulta validato per la traduzione inglese / italiano. La stessa procedura è operata dai parlanti madrelingua inglese sulle frasi derivate dal corpus italiano e raggruppate nello stesso Tipo. Si noti che non è necessario che lo stesso verbo italiano sia applicabile per la classe di azioni del Tipo 3 “John bends over”, che in italiano richiede appunto il riflessivo (*piegarsi*). La relazione *bend* / *piegare* non è quindi valida

“across-types”, ma è valida, come era auspicabile, “within-type”. In caso di mancata applicabilità la procedura richiede l’inserimento dei lemmi corretti e/o delle possibilità alternative (in questo caso *incurvare*) e si configura quindi come un sistema di induzione sia di tipi potenziali di azioni sia di verbi rilevanti.

Figura 69: L’interfaccia di validazione della produttività dei tipi CBE-Heavy



Tale valutazione è stata effettuata rispettivamente da parlanti madrelingua italiani e inglesi attraverso l’interfaccia presentata in figura 69 (interfaccia di CBE-Heavy) su tutte le occorrenze raggruppate nei tipi di IMAGACT (circa 50.000 frasi) e assicura una validazione pressoché totale delle entrate di IMAGACT ai fini della traduzione dei verbi di azione tra queste due lingue. In Moneglia et al. (2014a) sono riassunti i risultati di tale validazione. I validatori madrelingua inglese hanno giudicato che i verbi inglesi previsti per i tipi azionali di IMAGACT risultano essere una traduzione valida per le frasi italiane che costituiscono l’istanza di tali tipi nel 98,5 % dei casi e viceversa gli annotatori madrelingua italiani hanno giudicato il verbo italiano previsto per i tipi in IMAGACT adeguato a tradurre le frasi inglesi che ne costituiscono le istanze nel 98,87 % dei casi.

Tabella 5: Sommario della validazione della produttività dei tipi

	Verbs	Instances	Basic Types	Validated Instances
IT > EN	521	21.483	1.100	98,50 %
EN > IT	550	15.474	1.180	98,87 %

Lo stesso metodo è stato usato anche per evidenziare le necessità di implementazione dell'ontologia per il trattamento di concetti azionali lingua specifici e in particolare con riguardo alla lingua cinese. I casi d'uso trattati nel capitolo 5 sono conseguenza del lavoro di confronto di dettaglio della produttività in italiano di ciascun concetto in IMAGACT con la produttività dello stesso concetto in cinese. Tutte le volte che un concetto ha estensioni testimoniate da frasi italiane che risultano non traducibili da uno stesso verbo cinese, come nel caso delle estensioni di pulire o di abbassare viste in precedenza, si pone il problema del motivo e l'ontologia deve essere implementata per trattare i concetti specifici di quella lingua che sono necessari per il suo trattamento. La validazione compiuta sul cinese ha comunque mostrato un ampio valore predittivo dei concetti in IMAGACT. Pan (ress) ha mostrato in particolare che su 946 tipi considerati, le frasi italiane raccolte sotto ogni Tipo erano tutte traducibili da uno stesso verbo cinese nell' 84,14% dei tipi (796) e che i tipi su cui si concentra la necessità di categorizzazioni linguistiche diverse per trattare i concetti riferiti dalle frasi italiane assegnate al Tipo sono 150 (15,86% dei tipi complessivi in IMAGACT). In altri termini l'ontologia risulta predittiva per la categorizzazione delle azioni comuni, nella sua gran parte, anche dal punto di vista del cinese e le differenze di prospettiva linguistica si concentrano su un numero relativamente ridotto di azioni, studiate in dettaglio attraverso la stessa interfaccia di validazione. La seguente è una schermata dell'interfaccia di validazione che riporta un Tipo non validato. *suonare*, per un italiano è un verbo che si applica a tutte le attività musicali con gli strumenti e copre tutte le istanze di una ampia famiglia di prototipi. Non è così per un cinese, che non ha un verbo che copre tutti questi diversi tipi di eventi, come abbiamo visto nel paragrafo 5.2.1. La validazione porta in questo caso ad una implementazione della ontologia, in modo coerente con le linee di sviluppo suggerite dal modello insiemistico (il Tipo italiano generale di *suonare* corrisponde ad una famiglia in intersezione con tipi diversi del cinese).

8.1.1 Validazioni rispetto ai sistemi di traduzione automatica e disambiguazione

Sulla base di tali premesse, le validazioni compiute all'interno di questo progetto di tesi hanno come finalità di mostrare che IMAGACT, attraverso la suddivisione delle

interpretazioni di un verbo di azione in più Tipi azionali, può costituire una risorsa di riferimento migliore per task di Word Sense Discrimination rispetto alle ontologie correnti nello stato dell'arte delle ontologie che possono trattare i verbi delle lingue naturali, in particolare BabelNet e per quanto riguarda la traduzione automatica Google Translate. Sono stati condotti tre test di validazione i cui risultati possono essere letti a cascata.

L'obiettivo del primo test, riportato estensivamente in 8.3, è confrontare le possibilità teoriche di disambiguazione offerte rispettivamente dai BabelSynset di BabelNet e dai Tipi di IMAGACT. Il test si configura come un confronto preliminare sul modo di rappresentare i significati su cui i verbi possono variare nelle due risorse e da esso emerge una differenza sostanziale tra il livello di discriminazione del senso operata da IMAGACT, in cui le tipologie azionali sono definite su base pragmatica e ben riconoscibili, e quella di BabelNet, in cui le tipologie sono create su basi definitorie. I BabelSynsets risultano: a) compatibili con più Tipi azionali diversi; b) non completamente separati tra loro. Nello specifico il test mostra come la mancata discriminazione dei significati impatti effettivamente sulle possibilità di disambiguazione e traduzione.

In secondo luogo sono stati concepiti due test che mostrano in concreto quanto la specifica separazione dei concetti azionali contenuta in IMAGACT possa impattare positivamente sui risultati ottenibili in task di traduzione automatica. Questa parte del lavoro ha una rilevanza specifica in quanto la correttezza della traduzione è un dato molto più oggettivabile della correttezza dei sensi, di cui costituisce un correlato, e vi si può costituire un *benchmark*. Per validare il contributo che la base di conoscenza IMAGACT può apportare in task di traduzione automatica e assistita rispetto alle principali infrastrutture disponibili attualmente (BabelNet e Google Translate) sono stati disegnati due esperimenti di validazione, che misurano rispettivamente: a) la possibilità teorica di tradurre correttamente frasi riferite ai concetti azionali; b) il contributo potenziale di IMAGACT rispetto alle attuali capacità di risorse di riferimento attualmente a disposizione per la traduzione real-time, in particolare di Google Translate e del tool di disambiguazione BabelFy.

A tal fine è stato costituito un dataset rappresentativo della variazione semantica dei verbi generali d'azione dell'italiano e dell'inglese sia per quanto riguarda l'impatto in

frequenza che per la potenziale ambiguità di significato. Il paragrafo 8.4 dà conto di questa risorsa di riferimento per la validazione. Su questa base sono stati condotti gli esperimenti. Nel primo, a cui è dedicato il paragrafo 8.5, sono valutate in assoluto le possibilità traduttive derivanti dalla discriminazione di significato operata da IMAGACT rispetto a BabelNet e Google Translate. Vi si mostra in particolare che IMAGACT è potenzialmente più ricco e capace di esprimere un numero maggiore di relazioni di traduzione in quanto semplicemente contiene un numero maggiore di lemmi riferiti all'azione tra loro in relazione di traduzione potenziale.

Per valutare il contributo potenziale di IMAGACT rispetto alle attuali performance di Google Translate e BabelFy sono stati sfruttati sia l'ampia varietà di concetti azionali disponibili in IMAGACT sia l'insieme vasto di occorrenze di questi tipi nelle frasi italiane e inglesi presenti nel corpus. Tali frasi, raccolte nel dataset di *testing*, sono state date in input a queste risorse e il grado di traduzione e disambiguazione corretta rispettivamente raggiunta dai sistemi è stata valutata da madrelingua. I risultati mostrano un notevole deficit nelle capacità di disambiguazione delle frasi riferite a concetti azionali ottenibile con BabelFy. Più in particolare è stato verificato il forte deficit traduttivo evidenziato da Google Translate specificamente nei casi che costituiscono i differenziali semantici tra i traduttori principali in italiano e inglese. Questa informazione, valore aggiunto specifico di IMAGACT, si dimostra quindi essenziale per lo sviluppo di sistemi robusti di traduzione automatica.

Infine sono stati rapportati i risultati relativi agli errori di traduzione ai limiti di rappresentazione lessicale delle risorse, evidenziando una forte correlazione tra gli errori e l'assenza di traduttori (in particolare in Google Translate) e la parallela copertura lessicale assicurata da IMAGACT in quei casi (intorno all'80%).

8.1.2 *La validazione dei dati in input e output dell'ontologia IMAGACT*

Le validazioni realizzate in Gagliardi (2014) hanno particolare rilievo in questa sede, in quanto assicurano in primo luogo la validità dei dati su cui l'ontologia si basa e in secondo luogo sulla significatività e usabilità dei suoi risultati ai fini di disambiguazione da parte di umani, ovvero in task di traduzione assistita. In questo paragrafo riporteremo

per esteso i risultati su questo lavoro.

La prima validazione è dedicata a provare che i parlanti competenti di una lingua possono discriminare le occorrenze in cui un verbo è usato in senso proprio ed è riferito ad una azione fisica rispetto alle occorrenze che si riferiscono ad estensioni metaforiche o fraseologiche. Tale validazione fonda la selezione dei dati rilevanti alla costituzione di una ontologia dell'azione in corpora di parlato, nei quali l'uso dei verbi d'azione è spesso metaforico e non costituisce quindi un riferimento valido per l'induzione dell'ontologia dell'azione.

Il secondo aspetto validato è relativo all'effettiva usabilità della rappresentazione dei concetti di azione attraverso immagini prototipiche per il riconoscimento, da parte degli utenti potenziali, dei significati in cui un verbo d'azione è utilizzato. Per entrambe le validazioni sono state utilizzate le metriche tradizionali in linguistica computazionale per misurare il grado di convergenza di soggetti diversi nell'esecuzione dello stesso task (Carletta 1986). Un alto livello di *inter-tagger agreement* tra annotatori è infatti considerato indice della bontà e della riproducibilità di un paradigma di annotazione. L'assunto di base è che i dati siano "attendibili" se due o più annotatori sono in accordo nell'assegnare una categoria all'item in analisi. Un alto *agreement*, ovvero un ampio consenso tra gli annotatori, viene assunto quindi a garanzia della precisione intrinseca del processo di classificazione (Warrens, 2010; Artstein and Poesio, 2008) A partire dall'articolo di Carletta (Carletta, 1996) la metrica che costituisce lo standard de facto per la validazione di dati categoriali prodotti dall'annotazione di più annotatori umani sono le misure "kappa-style", ovvero coefficienti statistici che considerano l'influenza del caso nell'accordo raggiungibile dagli annotatori in un task di classificazione. Infatti, sebbene l'indice più semplice sia ovviamente la percentuale di accordo, ovvero la percentuale di giudizi su cui due analisti concordano nell'esecuzione indipendente di un task, detto anche *Observed Agreement* (Scott, 1955), tale valore non tiene in considerazione il peso che può avere il caso nel determinare di tale accordo. Ci sono vari motivi invece per tenerne conto. In primo luogo, dati due diversi schemi di codifica per lo stesso fenomeno, quello con il minor numero di categorie registrerebbe sempre una più alta percentuale di *agreement*. Inoltre l'*agreement* è influenzato fortemente dalla "prevalenza", ovvero la maggior frequenza di una delle categorie nei dati sottoposti a verifica. Se la

valutazione di un accordo percentuale non tiene conto che le categorie più semplici e frequenti hanno maggiore probabilità di *agreement*, il disaccordo sulle categorie meno frequenti viene automaticamente sottovalutato (Artstein and Poesio, 2008). Il livello di accordo nell'espressione di giudizi categoriali deve perciò essere esplicitato nei termini di eccedenza rispetto all'accordo ottenibile casualmente, pena la mancanza di informatività e significatività dell'indice utilizzato. I principali coefficienti per l'*agreement* tra due annotatori in questa linea sono il π di Scott e la k di Cohen (Scott, 1955; Cohen, 1960) la α di Krippendorff (2004) e, quando si deve valutare l'*agreement* tra più annotatori, la *weighted kappa* (Cohen, 1968), multi- π (Fleiss, 1971), multi- k (Mark Davies, 1982). Per l'applicazione di tali misure sono state seguite linee guida sia per la predisposizione di test di *reliability* sia indicazioni per la presentazione dei risultati (Di Eugenio and Glass, 2004; Artstein and Poesio, 2008; Bayerl and Paul, 2011) e si rimanda a Gagliardi (2014) per gli algoritmi.

8.1.2.1 *Test di identificazione della variazione primaria e marcata*

Con il primo test di validazione Gagliardi ha identificato il livello di *agreement* raggiungibile da annotatori esperti nella categorizzazione di occorrenze verbali nella categorie PRIMARIO/MARCATO, e quindi la replicabilità del criterio operativo derivato da Wittgenstein (1954) per effettuare tale discriminazione:

Una istanza è primaria se un parlante competente è disposto ad indicarla a qualcuno che non conosce la lingua come un possibile prototipo per l'azione 'Queste e simili cose sono ciò che noi chiamiamo "x"'. È un uso marcato altrimenti.

A due annotatori è stato sottoposto un set concordanze (derivate dal corpus oggetto di analisi in IMAGACT) di verbi di azione riconducibili ad un'area semantica coesa (girare e i lemmi verbali localmente equivalenti nei tipi della sua variazione) per un totale di circa 1.000 entrate. La tabella di seguito riporta le quantità di concordanze processate in questo test set, corrispondente a circa il 2% delle entrate processate in IMAGACT.

Il task consiste in un esercizio sottoposto agli annotatori esperti che hanno lavorato allo sviluppo dell'ontologia. Il protocollo di annotazione prevede infatti che ciascun

Tabella 6: Lemmi del test di validazione della distinzione Primario / Marcato

Lemma	Occorrenze
girare	654
capovolgere	5
curvare	4
mescolare	16
rigirare	20
rivolgere	150
ruotare	19
svoltare	9
volgere	30
voltare	67
TOTALE	974

annotatore, dopo aver letto ed interpretato l'occorrenza verbale in contesto, attribuisca il tag primario o marcato, ovvero discrimini gli usi fisici ed azionali da quelli metaforici o fraseologici, scartando di fatto i secondi dal data set di frasi con riferimento all'azione da cui viene indotta l'ontologia. La procedura da cui risulta l'*agreement* o il *disagreement* sullo stesso task è svolta dagli annotatori autonomamente ed indipendentemente, utilizzando un foglio di calcolo simile a quello riportato di seguito che evidenzia il dato in input, ovvero le concordanze dei lemmi in considerazione nel corpus Italiano su cui si basa IMAGACT.

Di tale valutazione riportiamo qui i risultati relativi ai casi in cui entrambi gli annotatori hanno svolto il task, ovvero dopo che questi hanno assegnato una interpretazione ad ogni occorrenza. Sono eliminati i casi nei quali uno o l'altro abbiano considerato non comprensibile una occorrenza e non svolto il task. I coefficienti statistici scelti per l'analisi di questo task sono in particolare il coefficiente k di Cohen, utilizzato per calcolare l'accordo tra due annotatori, insieme alla percentuale semplice di accordo (A_0 senza e con le correzioni per l'eliminazione dell'effetto di prevalenza $2A_0 - 1$ e le misure ulteriori π di Scott e α di Krippendorff.

Il livello di *inter-rater agreement* complessivo gratifica ampiamente la *reliability* del data set su cui si basa IMAGACT. Come è possibile vedere dalla tabella riassuntiva di seguito, l'accordo supera la soglia 0.9 ed è quindi un dato in assoluto molto alto rispetto ai valori, in genere non così alti, ottenuti in questo Tipo di task di categorizzazione linguistica.

Figura 70: Foglio di concordanze per la valutazione dell'agreement

	A	B	C	D	E
33	girare	23	/ gia // gia // si // si mette qui // si fa girare // SIL avvicina il registratore a ADR lascio gia // gia // si // si // aeh si // # mh // u u a chela // # è gia //	girare	/ vai // a gia // gia // si // si // aeh si // # mh // u u a chela / // si // gia // ADR indica un * immagine del libro ana // a ne ghè / chello ? questo
34	girare	24	// mh // danno u um gallo // il gallo // in punta // a gia // si	girare	? ah // lui gira // si // quando + mimmo // dopo // mi // mi // palla //
35	girare	25	danno u um gallo // il gallo // in punta // a gia // si // si // girare ? ah // lui	girare	// si // quando + mimmo // dopo // mi // mi // palla // palla // la palla
36	girare	26	l * immagine di un mulino a vento questo che è ? un mulino // no // no // la ruota che	girare	// li // ? si // // con le pale // questo quaggiù in fondo // si // fondo ?
37	girare	27	con le pale // questo quaggiù in fondo // si // fondo ? si // an // an // girare //	girare	/ si // # xxx // ci // come questo // SIL indaca ad ADR un cavallino di plastica si
38	girare	28	/ questa // si // è trasformata anche lei nell * elefante // mh // indica l * immagine della giostra	girare	? si // si // // la giostra // a gita // eh // la giostra // // qui c * è tutti
39	girare	29	/ è la giostra / anche questa // guarda che giostra grossa // questi gli aeroplanini // sono // na	girare	// tutti intorno / guarda / così // e tutti i bambini sopra // tutti i mimmi // fante // fante // fante
40	girare	30	gia // anche questi	girano	per bene // SIL aiuta ADR a voltare le pagine a no // chetto no // oo fante // quale vuoi ?
41	girare	31	l * immagine di un gelato qui c * è il gelato // chicco // ADR volta le pagine del libro vieni // si	girare	gira ? no // casa mamma // ah // a casa della mamma //
42	girare	32	qui // no // no // / a pumbe // / a mh + pumpe // a mh + pumpe // a mh + pumpe // a mh + pumpe // a mh + pumpe	girare	? no // casa mamma // ah // a casa della mamma //
43	girare	33	proviamo insieme ? dai // no // no // vuoi fare tu da sola ? ta # eh // cosa vuoi fare ? questo ? guarda	girare	! guarda Eleonora // eh // no // attenta che si rompe eh // mettiamo giù le forbici ? no //
44	girare	34	che	girare	andiamo su
45	girare	35	cosa spingi ? cosa vuoi toccare ? eh ? è duro ? non riesci a spingere ? che cos * è questo ? è il	girare	? guarda // Eleonora // hai visto che gira la rotellina dentro ? dov * è ? ci sono i pesci su
46	girare	36	registratore ? hai visto che	girare	queste carte ? guarda
47	girare	37	Eleonora // hai visto che	girare	la rotellina dentro ? dov * è ? ci sono i pesci su queste carte ? guarda Eleonora ! ci sono i
48	girare	38	registratore ? hai visto che	girano	? si muovono // mh ? èee // è la penna ? si // ascolta // vuoi un foglio
49	girare	39	Eleonora // ha fatto bene // adesso	girare	con una scarpa si e una no // eh // Eli // e ciuccio // o ciuccio // li hai fatti tutti tu i
50	girare	40	un * ape // e questo ? che cos * è questo ? eh ? è il bruco ? è il bruco questo ? #	girare	// così // uh // hai visto che qua il bruco insieme ai suoi fratelli ? c * era il mondo // poi chi
			cesto con le albicocche // e qui ? cosa sono ? ahm // ahm ahm ahm // hai fame proprio / eh	girare	sono gli altri
			allora // vhs che	giriamo	// vediamo cosa c * è qua dietro ? uh // i // i // mo // guarda ! pianino // c * è un delfino che

Tabella 7: Risultati del test 1 di Gagliardi

lemma	A_0	k	π	$2A_0 - 1$	α
capovolgere	1.0	1.0	1.0	1.0	1.0
curvare	1.0	-	-	1.0	-
girare	0.9809	0.9354	0.9354	0.9618	0.9534
mescolare	1.0	1.0	1.0	1.0	1.0
rigirare	1.0	1.0	1.0	1.0	1.0
rivolgere	0.99	0.9280	0.9279	0.98	0.9283
ruotare	1.0	1.0	1.0	1.0	1.0
svoltare	1.0	1.0	1.0	1.0	1.0
volgere	1.0	-	-	1.0	-
voltare	0.9354	0.6309	0.6309	0.8708	0.6339
TOTALE	0.9797	0.9565	0.9565	0.913	0.9565

8.1.2.2 Test di riconoscimento dei tipi azionali

Con un secondo test Gagliardi ha inteso verificare il livello di *agreement* raggiungibile da annotatori esperti nell'assegnazione delle frasi standardizzate ai tipi azionali IMAGACT, ovvero dimostrare la solidità della tipizzazione operata sui lemmi verbali oggetto di studio, mostrando che questa corrisponde a capacità discriminatorie effettive dei parlanti di una lingua. A tal fine è stato creato un set di 169 frasi standardizzate a partire dai materiali annotati nel corso del primo test così distinte per lemma.

Tabella 8: Lemmi del test di validazione della distinzione Primario / Marcato

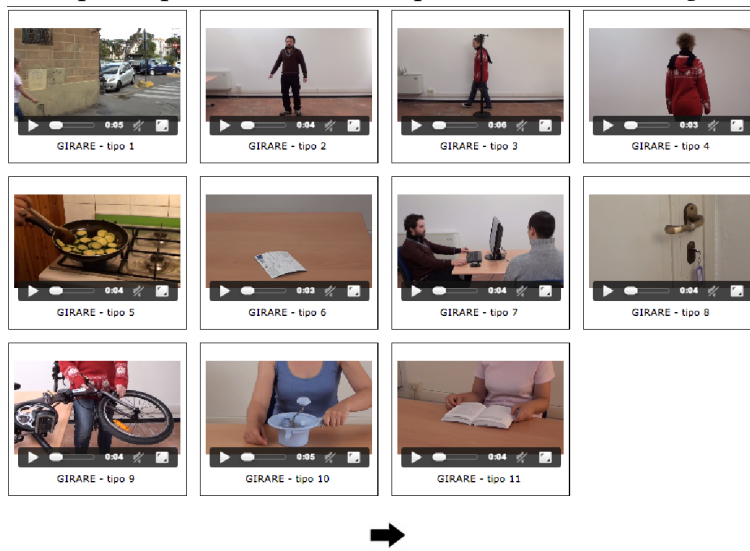
Lemma	Tipi	Occorrenze
girare	11	124
mescolare	4	16
rigirare	4	6
ruotare	4	8
voltare	3	15
TOTALE	26	169

A due mesi di distanza dal primo test, agli stessi annotatori, è stato chiesto di assegnare le frasi standardizzate di ciascun lemma ad un inventario dato di tipi, la variazione primaria identificata nel DB IMAGACT. La tipizzazione degli eventi è fornita agli annotatori mediante un ipertesto che consente di osservare l'insieme dei tipi azionali in IMAGACT e di avere a disposizione una schermata riassuntiva che agevoli la scelta.

Agli annotatori è stato richiesto di assegnare le frasi riportate nel foglio di seguito ad

8.1 IL QUADRO DELLA VALIDAZIONE DI IMAGACT

Figura 71: Iper testo per la presentazione dei tipi azionali del verbo *girare* agli annotatori



uno dei prototipi delle azioni estese dal verbo. La procedura, anche in questo caso, è stata svolta dagli annotatori autonomamente ed indipendentemente.

Di seguito sono sintetizzati i risultati secondo le stesse metriche del test precedente.

Tabella 9: Risultati del test 2 di Gagliardi

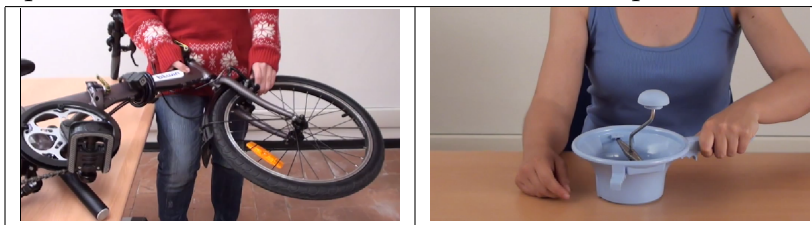
lemma	A_0	k	π	$2A_0 - 1$	α
girare	0.7903	0.7666	0.7649	0.5806	0.7659
mescolare	0.875	0.8	0.7981	0.75	0.8044
rigirare	1.0	1.0	1.0	1.0	1.0
ruotare	0.875	0.8297	0.8279	0.75	0.8387
voltare	1.0	1.0	1.0	1.0	1.0
TOTALE	0.8284	0.8167	0.8160	0.6568	0.8166

Il valore dei coefficienti, complessivamente superiore a 0.8, è considerato di eccellenza nella prassi di applicazione della k (e di sufficienza per la $k > 6$). In altri termini i due annotatori attribuiscono le standardizzazioni alle classi di azioni in modo sostanzialmente condiviso e pertanto è possibile dedurre che la tipizzazione ottenuta in IMAGACT dà luogo a giudizi altamente ripetibili. Il risultato è ancora più eclatante se si considera che, analizzando i dati meno positivi, relativi al verbo più generale (*girare*) si può vedere che il disaccordo (11 casi su 26) si concentra specificamente sulla coppia di tipi seguente.

La distinzione tra i due tipi è basata su un tratto semantico specifico: in entrambi i casi l'agente applica una forza sul tema, imprimendogli movimento rotatorio, ma nel Tipo 9 la

Figura 72: Foglio di standardizzazioni per la valutazione dell'*agreement*

	A	B
1	standardizzazione	numero del tipo
2	Il registratore gira	
3	Marco gira la pagina	
4	il giocattolo gira	
5	La ruota gira	
6	Il cavallino di plastica gira	
7	La ruota [della giostra] gira	
8	gli aeroplanini girano intorno	
9	La rotellina [del registratore] gira	
10	Marco gira il libro	
11	Fabio gira	
12	Fabio gira l'antenna	
13	Marco gira la rotellina [della radio]	
14	La trottola gira	
15	Fabio fa girare la trottola	
16	Fabio gira la rotella del volume	
17	Le rotelline girano	
18	Marco si gira	
19	Marco si gira [verso Cristina]	
20	Marco gira la chiave	
21	Fabio gira la tessera del puzzle	
22	Marco gira il pezzo del puzzle [per incastrarlo]	
23	Fabio si gira verso la finestra	

Figura 73: Tipi 9 (a sinistra) e 10 (a destra) della variazione primaria del lemma *girare*

forza è un impulso, mentre nel Tipo 10 la forza sull'oggetto è continua. La distinzione tra i due tipi di eventualità, chiara sul piano semantico, risulta probabilmente troppo granulare per consentire una decisione in un test con utenti privi di motivazione specifica, e i due concetti risultano indistinguibili agli annotatori. Ricalcolando i coefficienti unendo le due categorie in un concetto meno granulare si ottengono valori ancora più marcatamente eccellenti su ogni parametro ($A_0=0.879$, $k = 0.8606$, $\pi = 0.8605$ ed $\alpha = 0.8611$).

Il test 2 stato effettuato da Gagliardi anche per misurare il livello di *agreement* nell'assegnazione delle frasi ai tipi azionali anche da parte di annotatori non esperti. Questa ulteriore valutazione tende quindi a verificare la riconoscibilità e quindi la significatività della tassonomia azionale assegnata ad ogni verbo in IMAGACT anche per semplici

parlanti di madrelingua. L'esperimento testa lo scenario d'uso più semplice di IMAGACT, ovvero l'usabilità della risorsa in task di disambiguazione da parte di utenti generici. Questi, per poter tradurre, debbono disambiguare frasi naturali della propria lingua attraverso le scene di IMAGACT per prevedere quali verbi di una lingua diversa possono essere utilizzati per predicare di un Tipo specifico di azione istanziato da una frase. A tal fine sono stati reclutati quattro annotatori di livello culturale medio-alto non sottoposti ad alcun specifico training oltre la spiegazione del task. Anche in questo caso gli annotatori hanno lavorato in completa autonomia, seguendo lo stesso protocollo sperimentale dell'esercizio precedente. Per l'analisi dei dati sono state scelte le metriche per la valutazione della k tra più di due annotatori, in particolare la cosiddetta multi k di Davies e Fleiss (1982) e la multi- π di Fleiss (1971) oltre che A_0 e α di Krippendorff. Separatamente è stato evidenziato il calcolo dell'*agreement* su ogni copia (*Pairwise agreement*). Le tabelle di seguito sintetizzano i risultati.

Tabella 10: Risultati del test 3 di Gagliardi

lemma	A_0	multi- k	multi- π	α
girare	0.7258	0.6928	0.6923	0.6929
mescolare	0.8020	0.7025	0.6997	0.7046
rigirare	0.9166	0.88	0.8793	0.8844
ruotare	0.7708	0.6890	0.6871	0.6968
voltare	0.8222	0.6730	0.6652	0.6708
TOTALE	0.7504	0.7304	0.7301	0.7305

Tabella 11: Risultati del test 3 di Gagliardi: *Pairwise agreement*

lemma	C-D	C-E	C-F	D-E	D-F	E-F
girare	0.6122	0.8277	0.7564	0.6025	0.6119	0.7468
mescolare	0.6483	0.7018	0.7920	0.6483	0.5604	0.9006
rigirare	1.0	0.7692	1.0	0.7692	1.0	0.7692
ruotare	0.6666	0.6595	0.6521	0.8297	0.6666	0.6595
voltare	0.4791	0.8557	1.0	0.5833	0.4791	0.8557
TOTALE	0.6121	0.8277	0.7564	0.6061	0.6119	0.7468

In sintesi, i valori di agreement risultano intorno alla soglia di 0.7, e pur inferiori ai risultati ottenuti dagli annotatori esperti, sono comunque largamente accettabili per un task semantico da parte di non esperti.

Anche in questo caso, ad una analisi accurata dei dati, si può notare che il valore degli indici risulta più basso per il comportamento difforme di uno degli annotatori: se, sulla base dei risultati in tabella si selezionassero i migliori tre annotatori (C, E, F) i risultati non si discosterebbero dai valori ottenuti con gli esperti ($A_0 = 0,8224$, $\text{multi-}k = 0,8078$, $\text{multi-}\pi = 0,8077$, $\alpha = 0,8081$), mostrando così che, salvo giudizi idiosincratichi, sempre possibili in questo Tipo di compiti, i concetti espressi attraverso prototipi in IMAGACT corrispondono ad una categorizzazione dell'azione naturale e altamente riconoscibile. Questo risultato costituisce la premessa essenziale per l'utilizzo di IMAGACT per la disambiguazione da parte di umani ed è assunto anche per le analisi ulteriori condotte in questa tesi che utilizzano le scene di IMAGACT per la disambiguazione.

8.2 I CONCETTI AZIONALI IN BABELNET E IMAGACT

Il test seguente costituisce, come si diceva, un confronto tra i BabelSynset di BabelNet e i tipi di IMAGACT ed è finalizzato a verificare la funzionalità dei concetti definiti in tali ontologie nei task di disambiguazione. Per il test sono stati considerati i 5 verbi italiani a più alta frequenza: *mettere*, *prendere*, *passare*, *dare*, *portare*. Per ciascun verbo sono stati confrontati tutti i tipi azionali in IMAGACT con tutti i BabelSynset (BS) di BabelNet. Prima di tutto dobbiamo evidenziare una differenza cruciale tra queste due ontologie, ossia il trattamento della variazione marcata, che in IMAGACT è esclusa dalla documentazione della variazione del verbo, mentre BabelNet non opera una distinzione tra variazione primaria e marcata dei verbi. Pertanto, per poter fare una comparazione, abbiamo preliminarmente analizzato i BS e li abbiamo classificati in primari (P), marcati (M) e sia primari che marcati (PM), nel caso la definizione del concetto *f* comprendesse entrambe le variazioni. In BabelNet sono presenti infatti alcuni BS che fanno riferimento a concetti generali che si applicano sia ad oggetti fisici che a processi mentali. Un altro fattore da evidenziare è la presenza in BabelNet di BS provenienti da traduzione automatica che, almeno per i verbi analizzati in questo test, contengono per lo più rumore e comunque risultano inadeguati come "sensi" del lemma. Pertanto questi BS sono stati esclusi dalla comparazione. La seguente tabella mostra, per ogni verbo, il numero dei tipi di IMAGACT e il numero di BS di BabelNet, con una sotto-suddivisione che identifica

i BS Primari (P), Marcati (M), Primari-Marcati (PM) e quelli provenienti da traduzione automatica (MT), che sono stati separati dagli altri per escluderli dai conteggi.

Tabella 12: Confronto tra il numero di Tipi in IMAGACT e di BabelSynset in BabelNet per i 5 verbi italiani a più alta frequenza d'uso

	<i>mettere</i>	<i>prendere</i>	<i>portare</i>	<i>dare</i>	<i>passare</i>	TOTALE
Tipi IM	18	16	8	3	9	54
BS	10	37	18	19	16	100
P	4	6	3	3	2	18
M	5	18	11	11	13	58
PM	1	2	4	2	1	10
MT	0	11	0	3	0	14

Dalla tabella emerge una alta preminenza dei BS puramente marcati in BabelNet (più il 50%); parallelamente BabelNet rispetto a IMAGACT contiene un numero più limitato di concetti azionali: circa la metà, considerando il rapporto tra la somma P + PM (28) e i tipi di IMAGACT (54). La seguente tabella mostra che su questi 5 verbi ci sono 11 concetti azionali di IMAGACT che non sono presenti in BabelNet, mentre per nessun concetto primario di BabelNet manca la corrispondenza in IMAGACT (ultima riga). Questo dato evidenzia un deficit rappresentativo dell'estensione dei verbi d'azione in BabelNet, in cui alcuni concetti azionali non sono censiti nonostante che facciano parte della variazione primaria del verbo e siano dunque concetti produttivi nel linguaggio.

Tabella 13: Concetti azionali presenti in BabelNet e IMAGACT

	<i>mettere</i>	<i>prendere</i>	<i>portare</i>	<i>dare</i>	<i>passare</i>	TOTALE
Tipi di IM non in BN	5	2	1	1	2	11
BS non in IM	5	18	11	11	13	58
BS marcati	5	18	11	11	13	58
Concetti mancanti in IM	0	0	0	0	0	0

Analizzando i concetti azionali che sono rappresentati in entrambe le ontologie, vediamo che la categorizzazione dei BS è molto diversa da quella dei Tipi. Solo in pochi casi infatti c'è un matching tra un Tipo di IMAGACT e un BS (10 casi su 43, 23%); in 32 casi il Tipo è più specifico di un BS (più tipi mappano su un solo BS, 74%) e in un caso il Tipo è più generale del BS (più BS mappano su un solo Tipo).

Da questi numeri emerge chiaramente una differenza sostanziale nella partizione del

Tabella 14: Concetti più specifici, più generali e equivalenti tra BabelNet e IMAGACT

	<i>mettere</i>	<i>prendere</i>	<i>portare</i>	<i>dare</i>	<i>passare</i>	TOTALE
IM > BN	0	0	0	1	0	1
IM < BN	9	12	5	0	6	32
IM = BN	4	2	2	1	1	10

significato dei verbi d'azione nelle due ontologie; in particolare in BabelNet riscontriamo tre tipologie di problemi che possono avere un impatto negativo su disambiguazione e traduzione:

1. 1. Non tutti i concetti azionali sono censiti: su 54 Tipi di IMAGACT, 11 non trovano una rappresentazione in BabelNet (Tabella 13), cioè circa il 20%;
2. 2. In larga parte i BS rappresentano concetti più generali rispetto ai Tipi (Tabella 14): il 74% (32 casi su 43) dei concetti azionali che trovano rappresentazione in BabelNet sono sottodeterminati;
3. 3. Variazione primaria e marcata in BabelNet sono sovrapposte (Tabella 12).

Il 20% del punto 1 non significa che c'è una probabilità di 0,2 di non riuscire a disambiguare un'occorrenza di un verbo. Infatti prima di tutto questo valore è calcolato su una classe molto specifica di lemmi, i verbi d'azione ad alta frequenza, quindi qualsiasi induzione probabilistica deve essere circoscritta a questa categoria di verbi che, come abbiamo detto, sono particolarmente difficili da disambiguare. Inoltre la probabilità di occorrenza dei tipi è estremamente diversa tra Tipo e Tipo: in un verbo come mettere il Tipo più frequente occorre nel corpus 800 volte, mentre il meno frequente ha soltanto 4 occorrenze. I dati di questo test pertanto non possono darci una stima di quanto i concetti azionali mancanti in BabelNet impattino su possibili task di disambiguazione. Il fatto però che dei concetti azionali che sono produttivi e categorizzati linguisticamente da verbi ad alta frequenza d'uso non siano presenti in BabelNet in misura rilevante ha inevitabilmente un impatto negativo sull'efficacia di algoritmi di NLP che sfruttano l'ontologia. Di seguito offriamo due esempi concreti.

ESEMPIO 1: ERRORE DI DISAMBIGUAZIONE Il Tipo 1 (Fabio passa il pennello sulla superficie) e il Tipo 5 (La sarta passa il nastro nelle asole) del verbo passare non sono

rappresentati in BabelNet. Ciò determina l'impossibilità che un algoritmo di disambiguazione che sfrutti BabelNet possa assegnare il BS corretto ad una qualsiasi occorrenza di *passare* appartenente a questi due tipi. Le figure seguenti mostrano il risultato dello strumento di disambiguazione BabelFy (Moro et al., 2014b), che si basa su BabelNet¹. Come si vede in entrambi i casi il verbo *passare* è assegnato allo stesso BS (*bn:00088932v*, *attraversare – passare – transitare*), identificato dalla glossa "Go across or through", che evidentemente non ha nessuna attinenza al il concetto azionale riferito dal verbo nelle due occorrenze.

Figura 74: Disambiguazione della frase "La sarta passa il nastro nelle asole della giacca" con BabelFy



Figura 75: Disambiguazione della frase "Il pittore passa il pennello sulla tela" con BabelFy



ESEMPIO 2: ERRORE DI TRADUZIONE Il fatto che la maggior parte dei Tipi di IMAGACT che trovano rappresentazione in BabelNet faccia riferimento ad un BS più generale evidenzia un certo grado di vaghezza: in questi casi infatti concetti azionali distinti

¹ I BestExample sono stati leggermente modificati per renderli il più possibile prototipici e non ambigui

vengono categorizzati all'interno di un unico BS, che non potrà pertanto discriminare con precisione le categorie di azioni. Questo causa anche un problema in traduzione: i verbi di un BS troppo generale non sono infatti tutti applicabili all'intera classe di azioni che esso rappresenta, ma soltanto ad una sua sottoparte. Il Tipo 1 (Marco prende la matita dal tavolo) e il Tipo 13 (Il portiere prende la palla al volo) di prendere mappano sullo stesso BS (*bn:00088762v, prendere – afferrare*), descritto dalla glossa “Afferrare con le mani”. Inserendo le due frasi in BabelFy otteniamo infatti l'identificazione dello stesso BS associato a prendere, come mostrano le figure seguenti.

Figura 76: Disambiguazione della frase “Lo scrittore prende la matita dal tavolo” con BabelFy



Figura 77: Disambiguazione della frase “Il portiere prende la palla al volo” con BabelFy



Il problema di disambiguazione qui diviene un problema di traduzione, poiché i due concetti azionali si predicano in modo diverso in inglese. Il verbo *to take* è il

traducente corretto per prendere in azioni relative al Tipo 1, mentre le azioni del Tipo 13 si traducono con *to catch*. Inoltre l'applicabilità dei due verbi inglesi è esclusiva nei due tipi: *catch* non è applicabile alle azioni del Tipo 1 e *take* non è applicabile alle azioni del Tipo 13. I verbi inglesi contenuti nel BS in questione sono *take* e *get hold of*, mentre non è presente *catch*. Questo causa una traduzione sempre errata di *prendere* quando è un'istanza del Tipo 13, poiché verrà proposto *take* come traducente, che però non è applicabile. Un'ulteriore criticità nella discriminazione dei sensi di parola in BabelNet è la sovrapposizione tra BabelSynset, in cui c'è una sottodeterminazione delle glosse. Sempre in merito all'esempio precedente, infatti, nell'inventario dei sensi collegati a prendere troviamo il BS (*bn:00094731v*, *prendere*), descritto dalla glossa "Entrare in possesso" che evidentemente sembra anch'esso appropriato a contenere le azioni dei Tipi 1 e 13.

Figura 78: Verbi inglesi e italiani collegati al BS "bn:00088762v"



Figura 79: Verbi inglesi e italiani collegati al BS "bn:00094731v"



In questo caso quindi c'è anche un problema di ambiguità causata dalla sovrapposizione dei concetti espressi da due BS diversi.

8.3 COSTRUZIONE DEL DATASET DI RIFERIMENTO PER LA VALIDAZIONE DI IMAGACT

Per validare il contributo che la base di conoscenza IMAGACT può apportare in task di traduzione automatica e assistita rispetto alle infrastrutture disponibili attualmente è stato costituito un dataset che rappresenta i verbi generali d'azione dell'italiano e dell'inglese sia per l'impatto in frequenza che per la potenziale ambiguità di significato. Su questa base sono stati condotti gli esperimenti, rispettivamente: a) sulla copertura dei

traducenti; e b) sul successo delle disambiguazioni. I test sono riportati nei paragrafi 8.5 e 8.6.

Il dataset realizzato per questi esperimenti è stato concepito per fornire una rappresentazione significativa del potenziale dell'ontologia IMAGACT e comprendente 30 verbi d'azione ad alta frequenza rispettivamente in italiano e inglese sui circa 600 per lingua considerati in IMAGACT. Sono stati selezionati in particolare i verbi a maggior impatto nell'uso della lingua dal punto di vista della frequenza e insieme altamente ambigui rispetto al loro riferimento azionale. Nello specifico è stato considerato l'insieme dei verbi annotati in IMAGACT e si è proceduti per esclusione, eliminando:

- i verbi con un numero di tipi inferiore a 5 (1-4), ovvero con minore ambiguità potenziale;
- i verbi con un numero di occorrenze inferiore a 30, ovvero con minore probabilità di occorrenza nell'uso;
- i verbi considerati a priori poco azionali (ad es. *hold, keep, lasciare, tenere*);
- i verbi con una forte similarità semantica con altri verbi già selezionati (es. *levare-togliere, pigliare-prendere*), per non introdurre ridondanze.

Da questa lista sono risultati 30 verbi italiani e 30 verbi inglesi, che sono riportati nelle Tabelle 16 e 17. I lemmi del dataset coprono circa il 30% delle occorrenze verbali presenti nei corpora italiano e inglese ed è quindi altamente significativo. La Tabella 15 riporta il dettaglio con il numero di occorrenze verbali e di Tipi totali presenti in IMAGACT (relativamente ai lemmi annotati) e il numero di occorrenze e di Tipi che fanno parte dei verbi del dataset.

Tabella 15: Rappresentatività del dataset sull'insieme del corpus annotato di IMAGACT

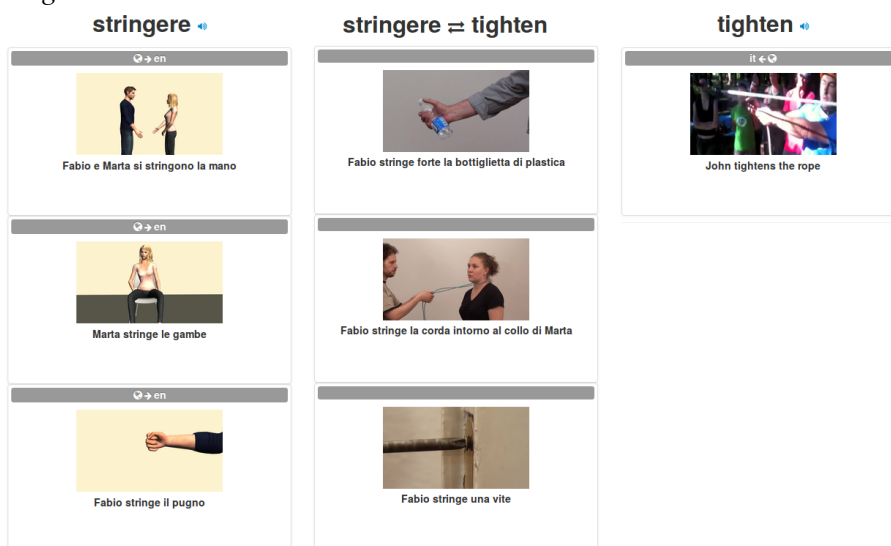
Lingua	Verbi annotati (VA)	Verbi dataset (VD)	Occorrenze VA	Occorrenze VD	Tipi VA	Tipi VD
inglese	803	30 (3,73%)	48.375	13.908 (28,75%)	1628	218 (13,39%)
italiano	695	30 (4,32%)	45.101	15.381 (34,10%)	1354	235 (17,35%)

Questi numeri evidenziano come la selezione dei lemmi verbali scelta con i criteri di cui sopra permetta di avere una buona copertura sia in termini di occorrenze (28,75% -

34,10%%), sia di Tipi azionali (13,39% - 17,35%), utilizzando soltanto un ridotto numero di verbi (3,73% - 4,32%).

Nella prospettiva di una valutazione dell'impatto per la traduzione le Tabelle 16 e 17 del dataset riportano per ogni verbo il corrispettivo traducevole tipico (TT) in italiano e in inglese, ossia il lemma che più spesso viene utilizzato nell'altra lingua per tradurre il verbo. Su questa base diviene possibile confrontare ogni verbo ed il suo TT: sono state contate le scene le scene collegate ad entrambi i verbi, rappresentanti i concetti azionali a cui è possibile riferirsi con uno dei due verbi, e suddivise tra scene comuni, in cui entrambi i verbi sono applicabili e scene differenziali, in cui uno solo dei lemmi è applicabile. Ad esempio in figura 80 è mostrato il differenziale tra il verbo italiano *stringere* ed il suo traducevole tipico inglese *tighten*.

Figura 80: Differenziale tra il verbo italiano *stringere* ed il suo traducevole tipico inglese *tighten*



Dalla comparazione dei concetti riferiti da questi due verbi notiamo che ci sono 7 scene totali ad essi collegate, delle quali 3 sono predicabili utilizzando entrambi i verbi, mentre 4 sono scene differenziali, in quanto è possibile utilizzare soltanto uno dei due verbi. Vediamo in particolare che *tighten*, a differenza di *stringere*, non è applicabile a parti del corpo, ma invece si estende ad un concetto azionale di "mettere in tensione una corda", che non è esprimibile con *stringere*.

Da questi numeri è stato estratto il valore PTT di un lemma, calcolato come il rapporto

8.4 VALUTAZIONE DELLA COPERTURA DEI TRADUCENTI

tra il numero di scene comuni con il traduceute tipico e il numero di scene totali collegate ad entrambi i lemmi; in questo esempio $PTT = 3/7 = 0,43$. Se assumiamo che ogni concetto azionale sia equiprobabile nell'uso linguistico, il valore PTT misura la probabilità che l'occorrenza di un lemma verbale sia traducibile con il suo traduceute tipico. Ovviamente le stime probabilistiche sull'incidenza delle tipologie azionali non sono significative al di fuori di un contesto specifico e quindi questa ipotesi di equiprobabilità dei concetti è solo teorica; nonostante ciò il valore PTT costituisce un indicatore efficace per capire quali sono i verbi che hanno maggiore o minore indeterminazione nella traduzione in un'altra lingua. Un valore PTT basso evidenzia un'alta indeterminazione nella traduzione di un verbo in lingua *source* verso la lingua *target*. Le tabelle (16 e 17) del dataset sono ordinate per PTT crescente.

8.4 VALUTAZIONE DELLA COPERTURA DEI TRADUCENTI IN IMAGACT, BABELNET E GOOGLE TRANSLATE

Per valutare la possibilità teorica di tradurre correttamente frasi riferite ai concetti azionali si è deciso di verificare il grado di copertura dei traduceuti in IMAGACT e nelle infrastrutture di confronto (BabelNet e Google Translate). Ciò determina il grado di completezza delle risorse per la traduzione automatica: in particolare, dato l'insieme di verbi nel data set di validazione, considerando la lista dei traduceuti inglesi di ogni verbo italiano (e viceversa), si verifica quante possibilità teoriche di traduzione in inglese vengono offerte da Google Translate (GT), da BabelNet (BN) e da IMAGACT (IM). In sintesi, questa analisi condotta sulla serie significativa di verbi d'azione identificata nel data set in 8.4 porta ad avere una misura della completezza lessicale delle diverse risorse per il loro sfruttamento in un task di traduzione. La completezza lessicale di una risorsa di traduzione è particolarmente rilevante nell'ambito della Computer Assisted Translation, in cui il sistema deve supportare il lavoro di un traduttore esperto. Qui le opzioni di traduzione possono essere fornite all'utente in modi diversi più o meno intelligenti e intuitivi, ma avere la lista completa delle possibilità lessicali di traduzione costituisce ovviamente un requisito fondamentale per l'efficacia di un sistema di questo Tipo anche per la traduzione automatica. Di seguito è descritta la procedura di valutazione della

copertura dei traducanti nelle tre infrastrutture dall'italiano all'inglese; la procedura inversa è analoga.

PROCEDURA DI VALUTAZIONE

1. Per ogni verbo italiano si compila la lista dei traducanti inglesi che sono:
 - a) per IMAGACT tutti gli equivalenti locali inglesi di tutti i tipi di quel verbo;
 - b) per BabelNet tutti i traducanti di tutti i BabelSynset "primari";
 - c) per Google Translate tutte le alternative di traduzione.
2. Un annotatore madrelingua inglese giudica se i verbi presenti nelle liste di BN e GT sono primari o marcati e, se sono marcati, li rimuove;
3. Un annotatore madrelingua italiano giudica se i traducanti inglesi proposti fanno parte di una variazione primaria o marcata del lemma italiano e, nel secondo caso, li rimuove.
4. L'insieme di elementi unici derivato dall'unione dei tre insiemi è la lista dei possibili traducanti del verbo che noi consideriamo completa.
5. L'annotatore controlla per ogni verbo della lista se esso è presente in IMAGACT, BabelNet e Google Translate.

Al termine di questa annotazione sono state estratte le statistiche relative alla copertura dei traducanti nelle tre risorse: per ogni risorsa viene calcolato l'indice di copertura come il rapporto dei traducanti disponibili in quella risorsa sul numero di traducanti dato dall'unione dei traducanti delle tre risorse. Le seguenti tabelle (Tab. 18 e 19) riportano una sintesi dei risultati del test; le tabelle complete l'insieme completo delle tabelle è invece riportato in Appendice 1.

Il dato bruto relativo alla percentuale di copertura sull'insieme dei traducanti teorici evidenzia la maggiore ricchezza di alternative lessicali in IMAGACT in entrambi i versi, sia dall'italiano all'inglese che viceversa. In particolare IMAGACT è significativamente più ricco rispetto a Google translate. Un'analisi di maggior dettaglio consente di evidenziare che l'intersezione dei traducanti, ovvero l'insieme delle possibilità traduttive comuni

nelle tre risorse, è scarsa sia dall'inglese all'italiano che dall'italiano all'inglese (tra l'8% e il 10 %). Questo fatto testimonia la scarsa standardizzazione delle relazioni lessicali nelle risorse per la traduzione. La maggior ricchezza di scelte traduttive in IMAGACT, ovvero la sua maggior ricchezza di dettaglio lessicale nel dominio del riferimento all'azione, è testimoniata dal fatto che sia per la traduzione in italiano che per la traduzione in inglese IMAGACT offre quasi il doppio delle scelte possibili.

D'altro lato, la completezza del repertorio di traducanti in IMAGACT è testimoniata dal fatto che il numero di verbi comuni sia a BabelNet che a Google Translate, ma assenti in IMAGACT, è ridotto (circa il 5% in entrambe le direzioni di traduzione). Nonostante ciò, è anche significativa la percentuale di traducanti assenti in IMAGACT ma presenti in una sola delle altre risorse (intorno al 30%); questi due dati danno a un diverso livello una misura dell'implementabilità di IMAGACT rispetto alle alternative di traduzione.

8.5 LA TRADUZIONE DEI CONCETTI AZIONALI IN BABELNET E GOOGLE TRANSLATE

Per valutare il contributo potenziale di IMAGACT rispetto alle attuali capacità delle risorse di riferimento a disposizione (Google Translate e BabelFy) abbiamo deciso di sfruttare l'ampia varietà di concetti azionali disponibili in IMAGACT e l'insieme vasto di occorrenze di questi tipi nelle frasi italiane e inglesi presenti nel corpus, come input per la traduzione attraverso queste risorse. Lo scopo di questa analisi è valutare l'efficacia delle comuni risorse di traduzione automatica nella traduzione delle frasi che predicano di un'azione e in parallelo valutare l'impatto che l'informazione presente in IMAGACT può avere nel momento in cui fosse masterizzata in sistemi analoghi.

Nello specifico, per ogni tipologia azionale di tutti i verbi presenti nel dataset sono state estratte più frasi diverse, tutte non considerate ambigue da un umano, ed è stata verificata la correttezza della traduzione, o comunque della disambiguazione del verbo da parte dei due sistemi. L'analisi condotta ha come scopo sia la misura generale dell'errore nella traduzione dei verbi d'azione nelle infrastrutture in oggetto, sia l'identificazione dei concetti che risultano più critici in traduzione, con particolare riferimento ai nuclei pragmatici in cui il traducante tipico di un verbo generale non è applicabile e che

costituiscono quindi il differenziale in estensione di due verbi in relazione di traduzione. Sono infatti questi ambiti, a cui l'informazione presente in IMAGACT è specificamente dedicata che necessitano della discriminazione fine del significato di un verbo in lingua *source* per poter effettuare una traduzione corretta in lingua *target*. A partire dal dataset di riferimento (Tabelle 16 e 17) sono state quindi estratte le occorrenze relative ad ogni Tipo e ed è stato eseguito un test di traduzione/disambiguazione utilizzando sia BN che GT secondo la procedura descritta di seguito.

PROCEDURA DI VALUTAZIONE Per ogni Tipo azionale sono state selezionate 3 standardizzazioni presenti in IMAGACT: le standardizzazioni sono state scelte utilizzando i seguenti criteri:

- nessuna ambiguità: le frasi selezionate sono tutte non ambigue per un parlante madrelingua, che quindi deve essere in grado di interpretarle univocamente;
- massimizzare la variazione orizzontale: per costituire un dataset sufficientemente vario sono state scartate frasi uguali o troppo simili ad altre già presenti dal punto di vista semantico;
- dare conto delle diverse realizzazioni sintattiche: laddove il verbo all'interno del Tipo può essere applicato con strutture sintattiche diverse (rappresentate da diversi BestExample), si è cercato di rappresentare questa varietà.

Per i Tipi con meno di tre testimonianze da corpus sono state aggiunte frasi semplici, utilizzando la competenza semantica dell'annotatore madrelingua, in modo da poter completare il test set. La lista di frasi inglesi ed italiane è stata tradotta sia con Google Translate che con BabelNet. Per GT è stato utilizzato Google Translate Toolkit², che produce un'unica traduzione della frase. Per BabelNet invece è stato utilizzato BabelFy, attraverso la API ufficiale (Moro et al., 2014a). BabelFy non è uno strumento nativo di traduzione, ma di disambiguazione; esso però è in grado di fornire una traduzione dei singoli lemmi della frase poiché identifica il concetto interlinguistico di ognuno di essi: pertanto i verbi in lingua *target* presenti nel BabelSynset disambiguato da BabelFy

² <https://translate.google.com/toolkit>

possono essere considerati tutti traduttori corretti del verbo in lingua *source* secondo questo sistema.

Per ogni frase tradotta un annotatore madrelingua giudica la correttezza del verbo selezionato dalla risorsa (GT o BN) per tradurre la frase di IMAGACT rispetto alla interpretazione definita dalle immagini del Tipo. Nel test non viene considerata la correttezza dell'intera frase tradotta, ma soltanto se il verbo in lingua *target* scelto per tradurre il verbo in lingua *source* è appropriato oppure no. La tabella completa con le frasi del dataset, i verbi utilizzati in traduzione da GT e BN ed i giudizi degli annotatori madrelingua è riportata in Appendice 2.

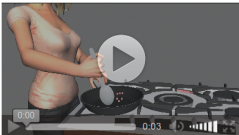


Le seguenti schermate, che rappresentano rispettivamente un Tipo azionale nella variazione di significato di girare in italiano e di turn in inglese, mostrano l'interfaccia di validazione realizzata e la procedura di validazione. L'interfaccia fornisce una terna di frasi per ciascun significato dei verbi in questione nel data set e sulla sinistra, i prototipi che specificano tali significati in IMAGACT, e quindi l'interpretazione da assegnare alle frasi. L'interfaccia permette l'esecuzione del task da parte di annotatori madrelingua della lingua *target* con conoscenza (non madrelingua) della lingua sorgente. L'annotatore valuta se il verbo selezionato, da Google e da BabelNet per la traduzione nella sua lingua è corretto, scorretto o inappropriato alla traduzione della frase della lingua sorgente, intesa secondo il modello di azione specificato nel prototipo. Se errato o inappropriato il verbo corretto è annotato. Lo stesso avviene nel caso che la lingua sorgente sia l'inglese.

Nei casi specifici le schermate esemplificano il task in due casi che ricadono nel differenziale di traduzione di due traduttori reciproci principali (*girare / turn*), ovvero significati di *girare* che non possono essere tradotti con *turn* e viceversa. Tipicamente, come vedremo nell'analisi dei risultati, i sistemi non sono in grado di selezionare un traduttore adeguato specificamente in questi casi.

DISCUSSIONE DEI RISULTATI La seguente tabella mostra i risultati generali della validazione sulla disambiguazione dei sensi dei verbi di azione compresi nel data set di validazione. Come era prevedibile, attraverso entrambi i sistemi oggetto di valutazione, risultati considerevolmente migliori si raggiungono quando la lingua *target* è l'inglese. Le performance di Google Translate sono però nettamente migliori rispetto a BabelNet

Figura 81: Interfaccia di annotazione per la valutazione delle traduzioni; *girare*.
girare (#641)

La cuoca gira il mestolo nella pentola
La mamma gira la pappa con la forchetta

Source	Google	BabelNet
 <p>Matteo gira il mestolo nella pentola</p>	<p>Matthew turns the ladle into the pot</p> <p>Verb is correct? <input type="radio"/> Yes <input type="radio"/> No <input checked="" type="radio"/> Unappr. <input type="radio"/> no verb</p> <p>Right Verb (N): <input type="text" value="move"/></p> <p>Main Verb (Y): <input type="text" value=""/></p>	<p>turn - ladle - pot</p> <p>Verb is correct? <input type="radio"/> Yes <input type="radio"/> No <input checked="" type="radio"/> Unappr. <input type="radio"/> no verb</p> <p>Right Verb: <input type="text" value="move"/></p> <p>Main Verb (Y): <input type="text" value=""/></p>
 <p>Cristina gira lo sciroppo con il cucchiaino</p>	<p>Cristina turns the syrup with spoon</p> <p>Verb is correct? <input type="radio"/> Yes <input type="radio"/> No <input checked="" type="radio"/> Unappr. <input type="radio"/> no verb</p> <p>Right Verb (N): <input type="text" value="stir"/></p> <p>Main Verb (Y): <input type="text" value=""/></p>	<p>Cristina - tablespoon</p> <p>Verb is correct? <input type="radio"/> Yes <input type="radio"/> No <input checked="" type="radio"/> Unappr. <input type="radio"/> no verb</p> <p>Right Verb: <input type="text" value="stir"/></p> <p>Main Verb (Y): <input type="text" value=""/></p>
 <p>La mamma gira le zucchine nella padella</p>	<p>Mom runs the zucchinis in the pan</p> <p>Verb is correct? <input type="radio"/> Yes <input checked="" type="radio"/> No <input type="radio"/> Unappr. <input type="radio"/> no verb</p> <p>Right Verb (N): <input type="text" value="stir"/></p> <p>Main Verb (Y): <input type="text" value=""/></p>	<p>ma - turn - frying_pan</p> <p>Verb is correct? <input type="radio"/> Yes <input checked="" type="radio"/> No <input type="radio"/> Unappr. <input type="radio"/> no verb</p> <p>Right Verb: <input type="text" value="stir"/></p> <p>Main Verb (Y): <input type="text" value=""/></p>

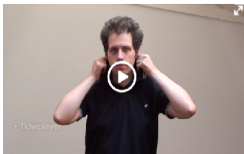
sia quando la lingua *target* è l'italiano sia quando è l'inglese. In particolare i risultati di disambiguazione con BabelNet risultano gravemente insoddisfacenti (49% di errore quando l'inglese è lingua *target* e ben 78% di errore se l'italiano è lingua *target*). Al contrario i risultati di Google Translate possono essere considerati apprezzabili (38% di errore con l'italiano lingua *target* e 21,5% di errore con l'inglese lingua *target*). Anche nel migliore dei casi, ovvero la traduzione in inglese dall'italiano con Google Translate, non è però possibile considerare robusti i sistemi di disambiguazione sottostanti a queste tecnologie quando l'oggetto è la disambiguazione di frasi riferite all'azione.

La strutturazione del data set però contiene non solo frasi, ma le raggruppa nelle diverse tipologie azionali che ogni verbo può estendere. Dall'analisi di dettaglio dei risultati ottenuti sul data set è possibile quindi osservare dove si addensa l'errore, se questo riguarda tipi specifici e in particolare se l'errore si addensa sui i tipi di azione che costituiscono differenziali tra traduttori principali. Per esempio, se data la relazione tra *girare* e *turn* come traduttori reciproci più probabili, l'errore si addensa sul loro differenziale: le azioni che sono *girare* ma non *turn* e viceversa.

La tabella seguente identifica nella parte alta il numero complessivo dei tipi di azione identificati nel dataset e, nella seconda parte in basso, i tipi di azione che costituiscono

Figura 82: Interfaccia di annotazione per la valutazione delle traduzioni; *turn*.
turn (#941)

John turns his collar up

Source	Google	BabelNet
 John turns his collar up	John gira il suo colletto Verb is correct? <input type="radio"/> Yes <input checked="" type="radio"/> No <input type="radio"/> Unappr. <input type="radio"/> no verb Right Verb (N): <input type="text" value="rigirare"/> Main Verb (Y): <input type="text"/>	meretricio - girare - bavero Verb is correct? <input type="radio"/> Yes <input checked="" type="radio"/> No <input type="radio"/> Unappr. <input type="radio"/> no verb Right Verb: <input type="text" value="rigirare"/> Main Verb (Y): <input type="text"/>
John turns his sleeve up	John gira la manica Verb is correct? <input type="radio"/> Yes <input checked="" type="radio"/> No <input type="radio"/> Unappr. <input type="radio"/> no verb Right Verb (N): <input type="text" value="alzare"/> Main Verb (Y): <input type="text"/>	Re_Giovanni - girare - manica Verb is correct? <input type="radio"/> Yes <input checked="" type="radio"/> No <input type="radio"/> Unappr. <input type="radio"/> no verb Right Verb: <input type="text" value="alzare"/> Main Verb (Y): <input type="text"/>
John turns the edge of the pants up	John trasforma il bordo dei pantaloni fino Verb is correct? <input type="radio"/> Yes <input checked="" type="radio"/> No <input type="radio"/> Unappr. <input type="radio"/> no verb Right Verb (N): <input type="text" value="alzare"/> Main Verb (Y): <input type="text"/>	girare - bordo - calzone Verb is correct? <input type="radio"/> Yes <input checked="" type="radio"/> No <input type="radio"/> Unappr. <input type="radio"/> no verb Right Verb: <input type="text" value="alzare"/> Main Verb (Y): <input type="text"/>

un differenziale tra il verbo della lingua sorgente e il traduttore principale nella lingua *target* e nell'ultima parte i tipi nei quali i traduttori principali trovano estensione comune.

In BabelNet l'addensamento dell'errore nelle frasi che istanziano tipi differenziali è drammatico: 98% nella traduzione in italiano della frase inglese e 82% nella traduzione della frase italiana in inglese. Sebbene anche da questo punto di vista i risultati di Google Translate siano nettamente migliori rispetto a BabelNet, l'incidenza dell'errore nei casi in cui il riferimento all'azione riguarda tipi differenziali risulta rilevante. Per quanto riguarda la traduzione in italiano delle frasi inglesi, mentre l'errore incide sul 56,42% dell'insieme dei tipi, l'errore sale all'81,67% dei tipi se si considera l'insieme dei tipi differenziali, ovvero quasi il doppio rispetto all'errore che si riscontra nei tipi co-estesi. Anche per la traduzione in inglese delle frasi italiane il dato della maggiore incidenza dell'errore specificamente sui tipi differenziali è analogo: mentre solo il 33,19% dell'insieme dei tipi tradotti contiene errore, la percentuale sale al 50% se si considera l'insieme dei tipi differenziali, ovvero quasi il doppio rispetto ai tipi co-estesi dai traduttori principali.

È possibile ragionare in modo più fine, considerando i casi in cui la traduzione di Google Translate si configura come un processo ripetibile, ovvero procede o meno correttamente su tutte le istanze de Tipo nel data set. Da questo punto di vista più qualitativo

possiamo vedere che, per quanto riguarda la traduzione dall'inglese all'italiano, la percentuale di traduzione pienamente corretta delle tre frasi che istanziano ogni Tipo, riguarda il 43,58% dell'insieme dei tipi, ma tale percentuale scende al 18,33 se si considerano i tipi differenziali. Tali percentuali scendono dal 66,81% al 50% nelle traduzioni dall'italiano in inglese. Significativamente l'errore procede tendenzialmente in modo sporadico sui tipi non differenziali (circa 8% EN>IT e 4% IT>EN dei tipi con tutte le frasi errate), mentre è assai più sistematico nei tipi differenziali (circa 55% EN>IT e 22% IT>EN dei tipi con tutte le frasi errate).

In sintesi il test di traduzione e disambiguazione con Google Translate e con BabelNet compiuto sulle frasi che istanziano i tipi di azioni in IMAGACT, che individuano le possibili ambiguità di un campione significativo dei verbi di azione italiani e inglesi, mostrano: a) l'inadeguatezza di BabelNet per la disambiguazione delle frasi che si riferiscono ad azioni; b) che una parte consistente degli errori di traduzione del verbo di Google Translate portano specificamente sui concetti azionali che costituiscono il differenziale semantico tra i traduttori principali. L'informazione specifica contenuta nell'ontologia dell'azione IMAGACT si configura quindi come una base di conoscenza che, se incorporata dai sistemi di disambiguazione e traduzione automatica, può effettivamente contribuire a rendere robuste tali tecnologie, costituendo un valore aggiunto per la traduzione automatica in un dominio, il riferimento all'azione, essenziale per il *processing* del linguaggio naturale.

8.6 RELAZIONI TRA I RISULTATI DEL TEST SUI TRADUCENTI CON IL TEST DI TRADUZIONE / DISAMBIGUAZIONE

È possibile integrare i risultati raggiunti nelle due valutazioni del confronto tra IMAGACT e Google Translate / BabelNet. In particolare diviene possibile controllare se nei casi di errore in traduzione la risorsa che commette l'errore ha nel suo inventario il traduttore corretto o meno. Questa informazione è già disponibile perché, in fase di validazione, il validatore ha sempre assegnato il verbo corretto alle frasi in cui era stato riscontrato un errore. Diviene quindi possibile evidenziare sia se il verbo che traduce correttamente le frasi errate è o meno un traduttore inserito nel sistema, sia se lo stesso traduttore

corretto è presente in IMAGACT. La tabella seguente illustra i risultati di questa verifica.

Il possibile impatto della mancanza dei traduttori teorici sull'errore risulta significativo, anche se in modo diverso nelle risorse a confronto. La minor ricchezza lessicale di Google Translate si riflette sul fatto che quasi la metà (47,15%) dell'errore dall'inglese in italiano correla con la mancanza del traduttore corretto. Questa percentuale è comunque alta anche dall'italiano all'inglese (circa 37%). Ovvero, in quei casi di errore, la traduzione corretta non poteva esserci a priori perché manca il verbo necessario. Se si considerano i dati percentuali, l'impatto della mancanza dei traduttori sull'errore è assai minore in BabelNet (poco più di un errore su 4), il che significa che i massivi problemi di disambiguazione riscontrati in questa infrastruttura dipendono semmai dalla sua struttura, ovvero dal sistema di disambiguazione e dal modo con cui sono definiti i sensi. Al contrario l'impatto potenziale della ricchezza lessicale con cui i diversi tipi di azione sono identificati in IMAGACT sulla correzione dell'errore risulta notevole. Sia per gli errori prodotti da Google sia per gli errori prodotti da BabelNet nelle traduzioni dall'inglese in italiano, nell'80 % dei casi circa in cui il verbo è assente nella risorsa risulta invece presente in IMAGACT. In entrambi i sistemi, nelle traduzioni dall'italiano all'inglese, intorno al 65% delle frasi errate prive di traduttore hanno un traduttore in IMAGACT.

In sintesi, l'identificazione fine dei tipi di azione sulle quali i traduttori delle due lingue divergono si accompagna in IMAGACT alla identificazione di traduttori specifici nei vari campi di applicazione del verbo, ovvero i verbi che possono essere localmente equivalenti nella variazione di un verbo generale su molti tipi di azione diversi. Entrambe queste informazioni si dimostrano altamente significative in relazione al Tipo di errori che gli attuali dei sistemi compiono nella traduzione e nella disambiguazione delle frasi riferite alle azioni, comprovando indirettamente il contributo che l'informazione contenuta in IMAGACT può fornire per rendere robusti sistemi che attualmente non garantiscono un trattamento corretto del linguaggio naturale nel modo con cui questo si riferisce alle azioni.

8.6 RELAZIONI TRA I RISULTATI DEI TEST PRECEDENTI

Tabella 16: Dataset italiano per i test di validazione

Verbo	Tipi	Occorrenze	Trad. Tipico	PTT
bloccare	5	84	stop	0,18
muovere	5	195	move	0,26
spostare	6	190	move	0,30
buttare	7	541	throw	0,35
separare	5	37	separate	0,36
piegare	6	47	bend	0,38
schiacciare	5	37	smash	0,40
tirare	7	546	pull	0,41
prendere	16	2459	take	0,42
stringere	6	56	tighten	0,43
appoggiare	9	118	lay	0,44
abbassare	6	52	lower	0,47
salire	5	213	climb	0,50
attaccare	8	206	attach	0,52
raccogliere	5	165	gather	0,55
allungare	5	37	extend	0,55
scendere	5	315	descend	0,56
spingere	12	166	push	0,59
sistemare	5	58	arrange	0,60
infilare	14	126	insert	0,62
togliere	9	300	remove	0,66
girare	11	654	turn	0,68
chiudere	9	495	close	0,70
alzare	9	349	raise	0,70
passare	6	946	pass	0,73
mettere	16	4018	put	0,74
portare	8	1509	bring	0,85
rompere	7	557	break	0,87
aprire	9	750	open	0,88
dividere	9	155	divide	0,91

8.6 RELAZIONI TRA I RISULTATI DEI TEST PRECEDENTI

Tabella 17: Dataset inglese per i test di validazione

Verbo	Tipi	Occorrenze	Trad. Tipico	PTT
knock	5	125	picchiare	0,13
set	5	434	mettere	0,19
drop	5	154	buttare	0,20
spread	8	70	sparpagliare	0,25
roll	6	48	rotolare	0,30
move	7	1004	spostare	0,30
collect	5	160	raccogliere	0,33
surround	5	45	circondare	0,38
carry	5	489	portare	0,39
hang	7	231	appendere	0,40
pull	17	220	tirare	0,41
take	13	4006	prendere	0,42
lay	6	126	appoggiare	0,44
press	7	122	premere	0,45
attach	5	54	attaccare	0,52
extend	6	69	allungare	0,55
play	7	615	giocare	0,56
fly	5	144	volare	0,57
push	18	143	spingere	0,59
rub	5	38	sfregare	0,60
lift	6	65	sollevare	0,63
remove	5	95	togliere	0,66
turn	10	702	girare	0,68
close	6	222	chiudere	0,70
raise	5	241	alzare	0,70
pass	7	339	passare	0,73
put	13	3231	mettere	0,74
break	6	307	rompere	0,87
open	7	258	aprire	0,88
divide	6	151	dividere	0,91

8.6 RELAZIONI TRA I RISULTATI DEI TEST PRECEDENTI

Tabella 18: Copertura traducanti in IM, GT e BN dall'italiano all'inglese

Verbo	IM	GT	BN	IM \cup GT \cup BN
mettere	87 (93,55%)	10 (10,75%)	12 (12,90%)	93
buttare	14 (77,78%)	5 (27,78%)	5 (27,78%)	18
spostare	56 (76,71%)	6 (8,22%)	20 (27,40%)	73
appoggiare	19 (76,00%)	6 (24,00%)	12 (48,00%)	25
togliere	31 (75,61%)	13 (31,71%)	10 (24,39%)	41
muovere	30 (75,00%)	7 (17,50%)	8 (20,00%)	40
attaccare	34 (69,39%)	16 (32,65%)	19 (38,78%)	49
allungare	9 (69,23%)	8 (61,54%)	7 (53,85%)	13
girare	18 (66,67%)	14 (51,85%)	10 (37,04%)	27
sistemare	32 (65,31%)	10 (20,41%)	17 (34,69%)	49
aprire	13 (65,00%)	6 (30,00%)	8 (40,00%)	20
rompere	19 (63,33%)	5 (16,67%)	12 (40,00%)	30
alzare	16 (61,54%)	11 (42,31%)	7 (26,92%)	26
prendere	46 (61,33%)	10 (13,33%)	41 (54,67%)	75
portare	16 (57,14%)	12 (42,86%)	16 (57,14%)	28
piegare	15 (55,56%)	12 (44,44%)	9 (33,33%)	27
spingere	13 (54,17%)	13 (54,17%)	8 (33,33%)	24
tirare	35 (53,85%)	17 (26,15%)	31 (47,69%)	65
salire	9 (47,37%)	7 (36,84%)	11 (57,89%)	19
dividere	12 (44,44%)	11 (40,74%)	14 (51,85%)	27
infilare	20 (44,44%)	13 (28,89%)	23 (51,11%)	45
schiacciare	17 (42,50%)	12 (30,00%)	25 (62,50%)	40
scendere	11 (40,74%)	12 (44,44%)	10 (37,04%)	27
abbassare	11 (37,93%)	11 (37,93%)	15 (51,72%)	29
passare	13 (37,14%)	16 (45,71%)	17 (48,57%)	35
chiudere	12 (36,36%)	9 (27,27%)	12 (36,36%)	33
separare	5 (35,71%)	4 (28,57%)	10 (71,43%)	14
bloccare	14 (33,33%)	8 (19,05%)	31 (73,81%)	42
stringere	11 (33,33%)	15 (45,45%)	24 (72,73%)	33
raccogliere	10 (30,30%)	13 (39,39%)	26 (78,79%)	33
SOMMA	648 (58,91%)	312 (28,36%)	470 (42,73%)	1.100
MEDIA	56,03%	32,69%	45,06%	

8.6 RELAZIONI TRA I RISULTATI DEI TEST PRECEDENTI

Tabella 19: Copertura traducenti in IM, GT e BN dall'inglese all'italiano

Verbo	IM	GT	BN	IM ∪ GT ∪ BN
put	86 (94,51%)	10 (10,99%)	7 (7,69%)	91
move	85 (92,39%)	9 (9,78%)	10 (10,87%)	92
knock	12 (92,31%)	7 (53,85%)	5 (38,46%)	13
attach	25 (89,29%)	6 (21,43%)	5 (17,86%)	28
hang	19 (86,36%)	5 (22,73%)	7 (31,82%)	22
surround	5 (83,33%)	2 (33,33%)	2 (33,33%)	6
rub	13 (81,25%)	5 (31,25%)	5 (31,25%)	16
push	17 (80,95%)	6 (28,57%)	3 (14,29%)	21
turn	23 (71,88%)	9 (28,13%)	16 (50,00%)	32
lift	28 (71,79%)	6 (15,38%)	13 (33,33%)	39
close	5 (71,43%)	3 (42,86%)	2 (28,57%)	7
pull	32 (71,11%)	10 (22,22%)	16 (35,56%)	45
remove	24 (70,59%)	12 (35,29%)	12 (35,29%)	34
press	9 (69,23%)	9 (69,23%)	8 (61,54%)	13
extend	13 (68,42%)	7 (36,84%)	14 (73,68%)	19
take	40 (65,57%)	15 (24,59%)	28 (45,90%)	61
collect	11 (61,11%)	6 (33,33%)	12 (66,67%)	18
lay	15 (60,00%)	12 (48,00%)	11 (44,00%)	25
set	26 (59,09%)	15 (34,09%)	20 (45,45%)	44
play	6 (54,55%)	6 (54,55%)	4 (36,36%)	11
open	10 (52,63%)	4 (21,05%)	9 (47,37%)	19
pass	12 (52,17%)	7 (30,43%)	18 (78,26%)	23
spread	12 (48,00%)	13 (52,00%)	20 (80,00%)	25
roll	11 (47,83%)	10 (43,48%)	12 (52,17%)	23
fly	7 (43,75%)	8 (50,00%)	7 (43,75%)	16
carry	5 (41,67%)	7 (58,33%)	4 (33,33%)	12
drop	17 (40,48%)	13 (30,95%)	23 (54,76%)	42
divide	6 (33,33%)	8 (44,44%)	14 (77,78%)	18
raise	7 (33,33%)	9 (42,86%)	12 (57,14%)	21
break	10 (20,00%)	13 (26,00%)	43 (86,00%)	50
SOMMA	591 (66,7%)	252 (28,44%)	362 (40,86%)	886
MEDIA	63,61%	35,20%	45,08%	

Tabella 20: Report di copertura dei traducenti

	EN > IT	IT > EN
Traducenti presenti in IM+GT+BN	95 (10,72%)	93 (8,45%)
Traducenti presenti in due risorse, escluso IM	48 (5,42%)	62 (5,64%)
Traducenti presenti solo in IM	415 (46,84%)	473 (43%)
Traducenti presenti in una sola risorsa ≠ IM	248 (27,99%)	379 (34,45%)
TRADUCENTI TOTALI	886	1100

8.6 RELAZIONI TRA I RISULTATI DEI TEST PRECEDENTI

Tabella 21: Errori di traduzione delle frasi

	EN > IT				IT > EN			
	GT		BN		GT		BN	
OK	400	61,92%	143	22,03%	507	78,48%	330	51,00%
NO	97	15,02%	152	23,42%	86	13,31%	171	26,43%
Unappr.	127	19,66%	96	14,79%	25	3,87%	48	7,42%
no verb	22	3,41%	258	39,75%	28	4,33%	98	15,15%
	646		649		646		647	

Tabella 22: Errori di traduzione nei diversi Tipi azionali

	EN > IT				IT > EN			
	GT		BN		GT		BN	
Tipi totali	218				235			
Tipi con errore	123	56,42%	194	88,99%	78	33,19%	148	62,98%
di cui in tutte le frasi (% su tot)	46	21,10%	146	66,97%	20	8,51%	69	29,36%
Tipi sempre corretti	95	43,58%	24	11,01%	157	66,81%	87	37,02%
Tipi differenziali	60				50			
Tipi con errore	49	81,67%	59	98,33%	25	50,00%	41	82,00%
di cui in tutte le frasi (% su tot)	32	53,33%	45	75,00%	11	22,00%	27	54,00%
Tipi sempre corretti	11	18,33%	1	1,67%	25	50,00%	9	18,00%
Tipi non differenziali	158				185			
Tipi con errore	74	46,84%	135	85,44%	53	28,65%	107	57,84%
di cui in tutte le frasi (% su tot)	14	8,86%	101	63,92%	9	4,86%	42	22,70%
Tipi sempre corretti	84	53,16%	23	14,56%	132	71,35%	78	42,16%

Tabella 23: Relazione tra l'errore di traduzione e la copertura dei traduttori

	EN > IT				IT > EN			
	GT		BN		GT		BN	
Totale traduzioni	653				637			
Totale traduzioni errate	246		506		139		317	
Verbo presente nella risorsa	130	52,85%	362	71,54%	87	62,59%	230	72,56%
Verbo assente dalla risorsa	116	47,15%	144	28,46%	52	37,41%	87	27,44%
di cui, verbo presente in IM	94	81,03%	116	80,56%	34	65,38%	59	67,82%
di cui, verbo assente in IM	22	18,97%	28	19,44%	18	34,62%	28	32,18%

VERSO UN'ONTOLOGIA WEB

In questo ultimo capitolo sono presentati dei lavori svolti su IMAGACT al fine di prefigurare la possibilità teorica di pubblicazione della risorsa in rete: in particolare è stata realizzata una versione dell'ontologia in formato OWL ed è stato effettuato il relativo mapping su *lemon*. Successivamente è stata condotta un'analisi per collegare l'ontologia con BabelNet, in modo da integrare la risorsa all'interno di uno strumento generale per la disambiguazione.

9.1 COSTITUZIONE DELL'ONTOLOGIA IMAGACT-OWL

Una delle ragioni che ha guidato la scelta di un database a grafo per la rappresentazione persistente dei dati strutturati in un modello insiemistico, è stata la vicinanza con un'ontologia RDF: infatti in questo formalismo basato su triple di Tipo soggetto-predicato-oggetto, le informazioni sono strutturate in un grafo orientato in cui i nodi sono collegati da relazioni. Una volta completata la migrazione di IMAGACT nel db Neo4J è stato quindi possibile convertire in modo semplice l'informazione in un formato ontologico: in particolare è stato utilizzato OWL, un'estensione di RDF e RDFS che permette una progettazione a più alto livello.

OWL ha tre componenti principali: le classi (Class), gli individui (Individuals) e le proprietà (Object e Data Properties). Gli individui sono gli oggetti presenti nell'ontologia; le classi sono strutturabili in gerarchia sulla base di una relazione *isA* ed ogni individuo appartiene a una classe; le proprietà sono relazioni tra individui (Object Property) o tra un individuo e un dato semplice (Data Property). Questa struttura di base è molto semplice, ma anche e molto espressiva: essa permette di modellare una rete semantica e di definire assiomi in grado di formulare inferenze tra gli oggetti presenti nell'ontologia.

Dei tre sottolinguaggi di OWL (OWL-Lite, OWL-DL e OWL-Full) è stato scelto OWL-

DL, che offre un'alta espressività, ma anche delle restrizioni sulle definizioni delle entità che consentono di produrre un'ontologia decidibile e completa. In questa prima versione sono state esportate sei classi dal database di IMAGACT, relative al nucleo informativo della struttura dati: Language, LanguageVariety, Verb, Type, Scene e Caption. Al fine di convertire i dati da Neo4J a OWL è stato implementato un algoritmo in Java che sfrutta la API REST di Neo4J, per estrarre i dati dal database attraverso query Cypher, e OWLAPI¹ per inserire i dati in un modello ontologico e produrre un'ontologia OWL.

Coerentemente con la struttura di nodi e relazioni di IMAGACT il modello dati in OWL prevede sei classi, una per ogni entità del database. Anche Object Property e Data Property rispecchiano abbastanza fedelmente lo schema di relazioni e proprietà presenti nel modello a grafo, ma con qualche differenza. L'entità Language assume il ruolo di lessico e LanguageVariety viene ad essere considerata una specializzazione di Language, in quanto rappresenta un sotto-lessico specifico di una varietà linguistica. Nel database a grafo, che implementa un modello insiemistico, avevamo una relazione di sottoinsieme: LanguageVariety \subseteq Language. Infatti, data una Language L ed una LanguageVariety LV tale che L HAS_VARIETY LV, allora

se un Verb V BELONGS_TO_VARIETY LV allora V BELONGS_TO_LANG L

Nell'ontologia, invece, la classe Language rappresenta il lessico comune a tutte le varietà e le LanguageVariety il lessico specifico. Ogni LanguageVariety eredita il lessico comune della lingua dalla classe Language.

LanguageVariety isA Language

Similmente anche le ObjectProperty belongsToLanguage e belongsToVariety sono state gerarchizzate, in quanto (utilizzando la terminologia precedente)

se V belongsToLanguage L allora V belongsToVariety LV

Le altre ObjectProperty hanno conservato il loro valore originale ed hanno la stessa semantica delle corrispondenti relazioni nel modello a grafo. Le proprietà dei nodi sono state portate in OWL come DataProperty e tipate utilizzando i DataType di OWL (*boolean*, *string*, *anyUri*, ecc.).

¹ <http://owlapi.sourceforge.net/>

Per l'esportazione sono state selezionate soltanto quelle più rilevanti, escludendo quelle derivanti dal processo di implementazione dell'ontologia; ad esempio sono state escluse tutte le informazioni riguardanti l'annotazione. Alla classe Language è stata infine aggiunta la DataProperty *code*, che contiene il codice della lingua secondo lo standard ISO-639; infatti le sigle utilizzate all'interno di IMAGACT non sono sempre conformi a questo standard.

L'ontologia OWL derivata ha permesso di ottenere una struttura adatta alla formulazione di query semantiche utilizzando il linguaggio standard SPARQL. Di seguito si riportano alcuni esempi di query.

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX im: <imagact-uri-da-definire#>
```

Elenco dei verbi in lingua inglese

```
SELECT ?v
WHERE
{
?v rdf:type im:Verb .
?v im:belongsToLanguage im:en
}
```

Elenco dei traducenti inglesi del verbo italiano prendere

```
SELECT distinct ?v
WHERE
{
?v rdf:type im:Verb .
?v im:belongsToLanguage im:en .
im:it_prendere im:refersTo ?s .
?v im:refersTo ?s
}
```

Elenco dei tipi di prendere

9.2 MAPPING SU *lemon*

```
SELECT distinct ?t
WHERE
{
  ?t rdf:type im:Type .
  im:it_prendere im:hasType ?t
}
```

Lista delle scene collegate ad un Tipo

```
SELECT ?s ?v
WHERE
{
  im:t_4009 im:containsScene ?s .
  ?s im:video_url ?v
}
```

Descrizioni testuali di una scena in una lingua

```
SELECT ?c ?t
WHERE
{
  im:s_7b37e13d im:hasCaption ?c .
  ?c im:captionVerb ?v .
  ?v im:belongsToLanguage im:it .
  ?c im:caption ?t
}
```

9.2 MAPPING SU *lemon*

In riferimento a quanto detto nel Capitolo 2, la rappresentazione delle informazioni attraverso le ontologie non è di per sé sufficiente alla costruzione della rete semantica che è alla base dei nuovi paradigmi del web. L'interconnessione delle informazioni e, di conseguenza, il *mapping* e il *linking* tra ontologie diverse divengono aspetti essenziali per l'accesso alla conoscenza e per il suo arricchimento, come testimoniato dagli sviluppi sempre maggiori della ricerca in questo ambito (Otero-Cerdeira et al., 2015). In generale le tecniche di mapping si basano sull'utilizzo di procedure automatiche o manuali che determinano la similarità tra i concetti definiti nelle diverse ontologie (Ehrig and Sure, 2004) e collegano tra loro gli oggetti più vicini. Il collegamento tra entità simili

delle diverse ontologie consente di massimizzare le connessioni tra risorse diverse, che possono essere accedute in modo unificato attraverso strumenti di interrogazione dedicati, come SPARQL. Per queste ragioni è stato realizzato un mapping sull'ontologia *top-level lemon* (Par. 2.1), che è oggi il più moderno standard per la pubblicazione di risorse linguistiche in rete. Il mapping tra le due risorse è stato effettuato utilizzando i costrutti `rdfs:subClassOf` e `owl:sameAs`. Di seguito si descrivono le scelte operate per integrare IMAGACT dentro la struttura di *lemon*.

La classe `Language` di IMAGACT, nella nuova concezione in OWL, è equivalente alla classe `Lexicon` di *lemon*: essa rappresenta un lessico di una lingua ed ha una `DataProperty` `language` che ne specifica la lingua in standard ISO-639. Pertanto sono state impostate le relative equivalenze tra classi e proprietà:

```
im:Language owl:sameAs lemon:Lexicon
```

```
im:code owl:sameAs lemon:language
```

La relazione gerarchica tra `Language` e `LanguageVariety` è stata preservata. *lemon* contiene una classe `LexicalEntry` che rappresenta un'entrata lessicale; la classe `Verb` è quindi stata mappata su `LexicalEntry`, ma come sottoclasse, dal momento che contiene soltanto una parte del lessico di una lingua: i verbi d'azione.

```
im:Verb rdfs:subClassOf lemon:LexicalEntry
```

Un ragionamento analogo è stato seguito per la classe `Type`, che rappresenta un Tipo azionale, e che è stata mappata come sottoclasse di `LexicalSense`: i tipi azionali costituiscono infatti un universo più specifico dei sensi di parola.

```
im>Type rdfs:subClassOf lemon:LexicalSense
```

In questa asserzione bisogna considerare che c'è uno slittamento di questa formalizzazione dalla concezione teorica di IMAGACT: ciò che qui viene asserito è che il Tipo azionale sia un senso di una parola, mentre in realtà esso rappresenta un campo di variazione primaria del senso. In questo contesto abbiamo ritenuto opportuno introdurre questa forzatura in favore di un pieno mapping con *lemon*; consideriamo che la potenzialità di IMAGACT come risorsa per task di Word Sense Disambiguation derivano proprio dal fatto che attribuiamo il ruolo di sensi di parola ai tipi azionali. Posto quindi che i `Type`

sono *LexicalSense*, non possiamo dire che queste due entità siano equivalenti: infatti la nozione di *LexicalSense* è generale, mentre il *Type* è specifico del lessico verbale azionale.

Ciò che non trova una possibilità di mapping all'interno di *lemon* è la classe *Scene*, dal momento che una scena non è un oggetto linguistico. Viceversa la *Caption*, che descrive la scena attraverso una frase semplice, è a tutti gli effetti un elemento pertinente ad una risorsa lessicale ed è stato mappato sulla classe *UsageExample*. Anche in questo caso è stata utilizzata *rdfs:subClassOf*, per specificare che una *Caption* non è un qualsiasi esempio d'uso di un lemma, ma è una specifica descrizione di un prototipo azionale.

`im:Caption rdfs:subClassOf lemon:UsageExample`

Vediamo adesso come sono state trattate le relazioni tra le classi, ossia le *ObjectProperty*.

L'aspetto più interessante di questo mapping è sicuramente la rappresentazione delle relazioni semantiche in *lemon*; all'interno della *ObjectProperty* *SenseRelation*, che esprime una relazione tra *LexicalSense*, troviamo infatti quattro sotto-proprietà: *Broader*, *Narrower*, *Equivalent* e *Incompatible*, che rappresentano rispettivamente le relazioni di maggiore o minore generalità, equivalenza e incompatibilità tra sensi. A queste abbiamo aggiunto una nuova *SenseRelation*, che è *Overlapping*, e rappresenta quei casi ampiamente discussi nella tesi in cui due Tipi non sono né più specifici, né più generali, né equivalenti, ma hanno un ambito di variazione comune.

Non solo IMAGACT rappresenta le relazioni semantiche che intercorrono tra Tipi, ma le deduce a partire dai dati e, in particolare, dai rapporti Tipo-Scena. Questo ha permesso di poter ridefinire queste 5 *SenseRelation* sulla base di regole logiche. Di seguito si riporta la formalizzazione di queste regole in sintassi logica.

$$\begin{aligned} \text{Narrower}(t_1, t_2) \Leftrightarrow & (\forall s_1 \in \text{Scene}, \text{containsScene}(t_1, s_1) \Rightarrow \text{containsScene}(t_2, s_1)) \\ & \wedge (\exists s_2 \in \text{Scene}, \text{containsScene}(t_2, s_2) \wedge \neg \text{containsScene}(t_1, s_2)) \end{aligned}$$

$$\text{Broader}(t_1, t_2) \Leftrightarrow \text{Narrower}(t_2, t_1)$$

$$\text{Equivalent}(t_1, t_2) \Leftrightarrow (\forall s \in \text{Scene}, \text{containsScene}(t_1, s) \Leftrightarrow \text{containsScene}(t_2, s))$$

$$\begin{aligned} \text{Overlapping}(t_1, t_2) \Leftrightarrow & (\exists s_1 \in \text{Scene}, \text{containsScene}(t_1, s_1) \Leftrightarrow \wedge \neg \text{containsScene}(t_2, s_1)) \\ & \wedge (\exists s_2 \in \text{Scene}, \text{containsScene}(t_2, s_2) \wedge \neg \text{containsScene}(t_1, s_2)) \\ & \wedge (\exists s_3 \in \text{Scene}, \text{containsScene}(t_1, s_3) \wedge \neg \text{containsScene}(t_2, s_3)) \end{aligned}$$

$$\text{Incompatible}(t_1, t_2) \Leftrightarrow (\forall s \in \text{Scene}, \text{containsScene}(t_1, s) \Rightarrow \neg \text{containsScene}(t_2, s))$$

La ObjectProperty Entry di *lemon* rappresenta il collegamento tra un Lexicon ed una LexicalEntry; questa è stata estesa creando la sottoclasse più specifica verbalEntry, che ha come Range Verb (che è una specializzazione di LexicalEntry). Rispetto alla definizione di belongsToLanguage, Entry rappresenta esattamente la proprietà inversa, quindi è stata annotata come inverse. Non c'è alcuna relazione diretta tra Type e Caption, ma è possibile sfruttare la proprietà Example di *lemon*, che collega un LexicalSense ad un UsageExample: ricordiamo infatti che ogni Type è un LexicalSense e ogni Caption è un UsageExample.

In questa nuova ontologia mappata è possibile eseguire query generali ad IMAGACT utilizzando le tassonomie di *lemon*. La figura seguente (Fig. 83) mostra un'istantanea dell'ontologia IMAGACT-lemon all'interno del software di gestione Protégé². Per l'accesso all'ontologia si veda l'Appendice 3.

9.3 COLLEGAMENTO IMAGACT-BABELNET

La tendenza crescente ad incrementare il mapping tra risorse deve fare i conti con il fatto che ogni ontologia è costruita con criteri differenti, che fanno capo a differenti quadri teorici; talvolta quindi non è possibile avere un sufficiente livello di similarità tra entità per poter effettuare un mapping. In questo contesto *l'instance matching* diventa particolarmente rilevante, poiché consente di collegare risorse in maniera "leggera", attraverso una connessione tra singole istanze delle ontologie (Castano et al., 2008; Nath et al., 2014). L'esperimento condotto si colloca nello scenario più generale in cui viene realizzato un collegamento tra un'ontologia di dominio aperto, che raccoglie un'ampia serie di concetti eterogenei poco specificati, e una specialistica, dove i concetti di un singolo dominio sono rappresentati in modo più dettagliato (Magnini and Speranza,

² <http://protege.stanford.edu/>

Figura 83: L'ontologia IMAGACT-lemon dentro Protégé

The screenshot shows the Protégé ontology editor interface. The top menu bar includes 'Active Ontology', 'Entities', 'Classes', 'Object Properties', 'Data Properties', 'Individuals by class', and 'SPARQL Query'. The main window is divided into three panes:

- Class hierarchy:** Shows a tree structure of classes. The 'Verb' class is selected and highlighted in blue. Other classes include 'Thing', 'Error1', 'Error2', 'Generates', 'Language', 'LanguageVariety', 'LemonElement', 'Argument', 'Component', 'Condition', 'Constituent', 'Context', 'Definition', 'Form', 'Frame', 'HasLanguage', 'Lexical entry', 'Lexicon', 'Morphological pattern', 'HasPattern', 'Lexical Condition', 'Lexical Context', 'Part of word', 'Phrase', 'Word', 'Lexical sense', 'Type', 'Lexical Topic', 'Lexicon', 'Morphological pattern', 'Morphological Transform', 'Node', 'Phrase element', 'Property Value', 'Prototype', 'Syntactic role marker', 'Usage Example', 'List', and 'Scene'.
- Usage: Verb:** Shows the usage of the 'Verb' class. It indicates 'Found 9494 uses of Verb' and lists several instances, each with a diamond icon and the text 'Type Verb'. The instances include:
 - ban_অনুকরণ_করা
 - ban_অভিনয়_করা
 - ban_অটকানো
 - ban_আকার_নেওয়া
 - ban_আটকানো
- Description: Verb:** Shows a list of descriptions for the 'Verb' class, including:
 - ban_আকার_নেওয়া
 - ban_আটকানো
 - ban_আটকে_রাখা
 - ban_আটকানো
 - ban_অনা
 - ban_অবৃত্তি_করা
 - ban_অলিঙ্গন_করা
 - ban_ওঠানো
 - ban_ওপ্তানো
 - ban_করাঘাত_করা
 - ban_কাটা
 - ban_কেলিত_করা
 - ban_খাওয়া
 - ban_খেলা_করা
 - ban_গেলা
 - ban_গোছানো
 - ban_গোটানো
 - ban_গোল_করা
 - ban_গড়ানো

2002). Il collegamento tra ontologie specialistiche e generaliste consente di avere un quadro ontologico comune a cui algoritmi diversi possono fare riferimento per identificare i sensi a diversi livelli di granularità.

L'analisi di seguito descritta riguarda un'ipotesi di collegamento tra due ontologie linguistiche, BabelNet e IMAGACT, entrambe multimediali, multilingui e sfruttabili per task di traduzione e disambiguazione (Moro and Navigli, 2015; Russo et al., 2013; Moneglia, 2014a). Il collegamento tra le ontologie avviene attraverso la componente visuale di IMAGACT, ovvero la rappresentazione delle azioni per mezzo di scene prototipiche.

9.3.1 *Tipi azionali e BabelSynset*

Da un punto di vista di un task di Word Sense Disambiguation i Tipi azionali di IMAGACT sono equiparabili ai BabelSynset, in quanto costituiscono partizioni del significato di un lemma in contesto interlinguistico. In realtà la concezione e le modalità costruttive delle due risorse non ci consentono di avere una sufficiente similarità tra Tipi e BS da permetterne un mapping. Un'evidenza pratica di questo la si può avere dalla tabella 14 dell'esperimento preliminare del paragrafo 8.2, in cui si vede quanto siano rari i BS ed i Tipi con un alto livello di similarità: nella maggior parte dei casi infatti i BS fanno riferimento a concetti molto generali e non univoci.

Una considerazione più teorica riguarda invece la diversa natura di BS e Tipi: un BS infatti è un aggregatore di *synset* (provenienti in massima parte da WordNet, ma anche da altre risorse) in lingue diverse. In questa categorizzazione vengono appiattite le differenze semantiche nelle varie lingue: si assume che *synset* diversi di lingue diverse facciano capo allo stesso concetto semantico universale. Questo è molto diverso dall'approccio di IMAGACT teso invece ad esplicitare in modo preciso le differenze di variazione dei lemmi nelle varie lingue sulla base di concetti pragmatici. Per quanto detto nel paragrafo 8.1, è vero che un buon numero di concetti azionali sono inter-linguisticamente produttivi anche tra due lingue lontane (come l'italiano e il cinese), ma questo numero è destinato ad abbassarsi progressivamente all'aumentare delle lingue e non possiamo fare assunzioni sull'esistenza di concetti universali. Vediamo quindi che, sia da un punto di vista teorico, sia da evidenze sperimentali, un mapping coerente tra BS e Tipi non è una strada percorribile.

9.3.2 *Test di linking tra le ontologie*

Il test sfrutta i verbi equivalenti in traduzione nelle due risorse per effettuare il *linking* tra le ontologie attraverso le scene prototipali presenti in IMAGACT. Il criterio utilizzato per trovare le scene che possono rappresentare i concetti azionali di BabelNet è quello della maggior corrispondenza tra i verbi collegati allo stesso BS e quelli collegati alle scene di IMAGACT. Per il test è stata utilizzata la versione 3.0 di BabelNet; i dati sono stati

Lingue	Verbi BN	Verbi IM
English	57.996	1.299
Spanish	16.832	735
Italian	15.590	1.100
Portuguese	11.517	792
German	5.210	992
Chinese	4.299	1.171
Norwegian	2.227	107
Danish	1.980	73
Polish	1.910	1.145
Hindi	342	189
Urdu	187	73
Bangla	117	120
Serbian	91	1.145
Sanskrit	35	198
Oriya	6	140
Totale	118.339	9.273

Tabella 24: Le 15 lingue comuni di BabelNet (BN) e IMAGACT (IM) con il relativo numero di verbi.

estratti utilizzando l'API Java (Navigli and Ponzetto, 2012b).

Un diverso tentativo di mapping su IMAGACT era stato precedentemente condotto da De Felice et al. (2014) attraverso l'analisi della collegabilità tra i tipi azionali di IMAGACT e i *synset* di WordNet (per l'inglese) e ItalWordNet (per l'italiano). Anche in quell'esperimento è stato sfruttato l'insieme dei verbi comuni come indice di similarità tra gli oggetti delle diverse ontologie. L'ambito di applicazione del nostro esperimento è però diverso, poiché è allargato ad un contesto multilingue: prima di tutto non si tratta di un mapping tra concetti delle due ontologie, ma di un *linking* tra istanze; in secondo luogo non si usa WordNet, ma BabelNet e, coerentemente, il dato di IMAGACT che viene sfruttato non corrisponde ai tipi (che dipendono dai diversi lemmi delle varie lingue) ma alle scene (oggetti di riferimento interlinguistico).

L'esperimento è stato condotto su un dataset di riferimento annotato manualmente e composto da 25 scene di IMAGACT e 30 BS; per ciascuna coppia <scena,BS> è stato giudicato se la scena fosse o meno adatta a rappresentare il BS. Il dataset, comprendente 750 giudizi binari (25 scene per 30 BS), è stato utilizzato per la valutazione degli algoritmi

automatici³. Per selezionare le scene sono stati considerati i prototipi azionali della variazione di 7 verbi inglesi (*put, move, take, insert, press, give e strike*), che proiettano 152 BS totali; di questi, ne sono stati scelti 25. La selezione è stata fatta in modo casuale, ma ha tenuto conto del fatto che non tutte le scene sono collegate allo stesso numero di verbi, sia perché le lingue di IMAGACT non sono egualmente rappresentate (vedi tabella 24), sia perché è diverso il numero di verbi utilizzabili per riferirsi alle azioni nelle diverse lingue. Per creare un test set che fosse un campionamento rappresentativo, abbiamo cercato di preservare queste differenze quantitative inserendo scene con un diverso numero di verbi collegati (da un minimo di 7 a un massimo di 18) in modo proporzionale all'intero set di scene di IMAGACT. Anche il numero di verbi contenuti in un BS è molto variabile, per cui la loro selezione ha seguito un criterio simile: ognuno dei 30 BS del dataset ha da un minimo di 4 ad un massimo di 51 verbi collegati. Il dataset pubblicato è derivato da un *agreement*: sono stati utilizzati tre annotatori ed il giudizio inserito è quello di almeno due annotatori su tre. L'*inter-rater agreement* riporta una *k* di Fleiss pari a 0,76.

L'algoritmo base utilizzato per il *linking* si avvale di una funzione che calcola la vicinanza tra una scena e un BS misurando la frequenza con cui i verbi collegati alla scena di IMAGACT sono legati anche al BS. L'insieme dei candidati è composto da tutti i BS che sono concetti possibili per ogni verbo collegato alla scena.

Algorithm 1 Algoritmo base

- 1: s : scena in ingresso
 - 2: V : set di verbi collegati a s in IMAGACT
 - 3: List(BabelSynset) LS : lista vuota
 - 4: per ogni v_i in V
 - 5: List(BabelSynset) Syn = lista di BS collegati a v_i
 - 6: aggiungi Syn a LS
 - 7: List(BabelSynset) FLS = $freqList(LS)$
 - 8: Collega s ai primi n BabelSynset di FLS
-

La funzione *freqList* calcola la lista di frequenza dei BS in LS e li ordina dal più frequente al meno frequente.

A partire da questa versione di base dell'algoritmo di *linking*, è stata implementata una versione migliorata che sfrutta la rete semantica di BabelNet, in modo da includere

³ Il dataset è disponibile in Appendice 4

nell'analisi anche i BS semanticamente vicini. Anziché estrarre gli equivalenti in traduzione soltanto dai BS collegati direttamente ai verbi, vengono considerati anche i verbi appartenenti a *synset* collegati al BS principale tramite le relazioni di BabelNet fino a un certo livello di profondità. Per pesare in modo differenziato i BS collegati al verbo abbiamo utilizzato una funzione ricorsiva w , definita nel modo seguente.

Dato S l'insieme dei BS, $s_0 \in S$ collegato direttamente al verbo e $s', s'' \in S$ collegati tra loro da una relazione $r \in R$, definiamo una funzione $w : S \rightarrow [0, 1]$ tale che $w(s_0) = freq(s_0)$ e $w(s'') = w(s') \cdot c \cdot p(r)$, dove: $freq$ calcola la frequenza del BS così come riportato nell'algoritmo base; R è l'insieme delle relazioni tra BS verbali; $p(rel) : R \rightarrow [0, 1]$ è una funzione che assegna un peso ad ogni relazione R ; $c \in [0, 1]$ è un coefficiente di riduzione di peso all'aumentare della distanza dal nodo centrale.

Questa metrica consente di differenziare l'importanza delle diverse relazioni semantiche di BabelNet. Abbiamo infatti verificato che, mentre alcune di esse sono molto rilevanti per il task, altre importano informazione non pertinente e devono quindi essere escluse. La tabella seguente mostra l'elenco delle relazioni tra BS verbali con il relativo valore di rilevanza misurato con *information gain* sul dataset annotato.

Relazioni in BabelNet	Valore IG
Hyponym	0.135
Also See	0.050
Hypernym	0.041
Verb Group	0.039
Entailment	0.009
Gloss Related	0.000
Antonym	0.000
Cause	0.000

Tabella 25: Relazioni tra BS verbali.

I due algoritmi sono stati eseguiti sulle 25 scene del dataset, quello di base e quello migliorato (considerando soltanto un livello di profondità). È stata quindi verificata l'aderenza dei primi n BS estratti dagli algoritmi alle scene. La tabella 26 riporta la sintesi dei risultati ottenuti (Per i risultati completi consultare l'Appendice 4).

Per entrambi gli algoritmi il primo BS candidato per il *linking* è sempre corretto. I risultati peggiorano progressivamente al crescere di n e, parallelamente, aumenta anche il

	Alg. base	Alg. migliorato
% corr. ($n = 1$)	100%	100%
% corr. ($n = 2$)	84%	88%
% corr. ($n = 3$)	76%	83%

Tabella 26: Percentuale di assegnazioni corrette di scene a BS con i due algoritmi e al variare di n .

divario qualitativo tra i due algoritmi. Inoltre, il dataset annotato è stato utilizzato come *training set* per due ulteriori algoritmi di *machine learning*, uno che considera soltanto *features* relative ai BS collegati in modo diretto, l'altro che include anche *features* dei BS collegati indirettamente attraverso relazioni semantiche. Il risultato riporta rispettivamente 81,16% e 86,95% di assegnamenti corretti di scene al BS. L'algoritmo utilizzato è SVM con kernel lineare e la valutazione è fatta con *10-folds cross-validation*. Benché il test set sia troppo piccolo per avere una stima precisa sull'efficacia, i risultati sono incoraggianti e compatibili con quelli ottenuti dai due algoritmi (semplice e migliorato) eseguiti sul dataset. La differenza tra le due percentuali mostra chiaramente che l'utilizzo dei BS vicini è significativo per questo task.

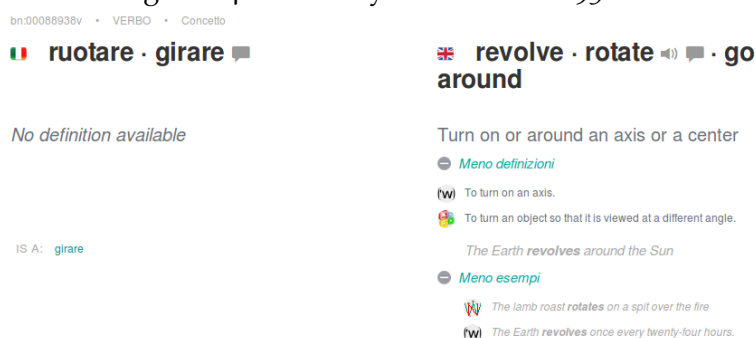
Benché non sia ancora stato fatto un *fine-tuning* dei parametri (per il quale è necessario un test set più ampio), i buoni risultati ottenuti da questo esperimento aprono alla possibilità di collegare le due ontologie attraverso le scene di IMAGACT, al fine di arricchire entrambe le risorse. Da un lato, i video di IMAGACT potrebbero rappresentare i concetti azionali dei BS; dall'altro, IMAGACT verrebbe arricchita con l'informazione di traduzione presente in BabelNet.

Dall'osservazione di BabelFy è apparso evidente che l'ipotesi di *linking* qui proposta avrebbe un notevole impatto sull'espressività della rappresentazione visuale delle frasi: BabelNet infatti contiene già un *repository* di immagini collegate ai BS nominali ed un collegamento tra le due ontologie può arricchire la rappresentazione dei BS verbali con i video 3D delle azioni presenti in IMAGACT. Oltre a questo, come mostrato nel paragrafo 8.2, gran parte dei concetti azionali di BabelNet sono sottospecificati e un *linking* con IMAGACT ne può fornire una identificazione maggiormente dettagliata. Un algoritmo di disambiguazione potrebbe quindi operare su due livelli, determinando in una prima fase il senso più generale di un verbo attraverso la scelta del BS più appropriato e

successivamente raffinando la selezione su un Tipo di IMAGACT. La scena diventa in questo modo un veicolo tra le due ontologie, per cui a partire da un BS è possibile ottenere tutte le scene ad esso collegate e tramite la scena è possibile (per ogni lingua) andare a vedere quali sono i tipi specifici che contengono quella scena.

ESEMPIO “bn:00088938v” è uno dei BS collegati al verbo italiano *girare*, è collegato a due verbi italiani (*ruotare* e *girare*) e a tre verbi inglesi (*revolve*, *rotate*, *go around*) ed è descritto dalla glossa inglese molto generale “Turn on or around an axis or a center” (Fig. 84).

Figura 84: Il BabelSynset “bn:00088938v”



Attraverso il collegamento con IMAGACT questo BS verrebbe ad essere rappresentato da 4 scene, corrispondenti a 4 diversi concetti azionali che rientrano tutti nell’area semantica coperta dal BS, coerentemente con la sua definizione. Ciascuna scena circoscrive un campo semantico più specifico ed è collegata a verbi diversi nelle diverse lingue (Fig. 85).

Figura 85: Collegamento con IMAGACT del BabelSynset “bn:00088938v”



CONCLUSIONI

In conclusione, i predicati verbali delle lingue naturali che si riferiscono all'azione sono altamente ambigui, perché possono riferirsi a tipi di azione diversi e lingue diverse segmentano in modo diverso questo universo primario dell'esperienza e dell'attività umana. La mancanza dell'informazione ontologica, sia sul *range* di variazione dei verbi di azione di una lingua sia sulle diverse modalità di riferimento all'azione dei verbi delle lingue, incide fortemente sulla possibilità di disambiguare e tradurre frasi semplici riferite alle azioni. Il capitolo 8 della tesi ha dato una misura obiettiva di questo deficit, che impatta in modo significativo sull'effettiva usabilità delle tecnologie di traduzione automatica e sugli strumenti di disambiguazione che, oltre al dominio della traduzione, sono necessari per la comunicazione naturale uomo-macchina.

IMAGACT, attraverso un metodo induttivo basato su corpora rappresentativi del riferimento all'azione in italiano e inglese, ha dato un contributo importante in questo senso creando una base di conoscenza che identifica le azioni a maggior impatto nella vita quotidiana, specificando al contempo le diverse relazioni che verbi italiani, inglesi, spagnoli e cinesi hanno con tale universo. Del *framework* di IMAGACT, che ha costituito il background del progetto di tesi si è dato conto nel terzo capitolo. Con questo lavoro abbiamo mostrato però che il modo con cui le lingue naturali si riferiscono all'azione introduce una complessità sia intra-linguistica che interlinguistica nella ontologia dei concetti naturali che non può essere trattata coerentemente con il consueto modello gerarchico delle relazioni tra entità ontologiche: le relazioni di equivalenza tra i diversi modi con cui ci si può riferire ad una azione (*equivalenza locale*) e la selezione esatta delle relazioni inferenziali necessitano modelli di rappresentazione diversi se si vuol dare una informazione consistente e empiricamente fondata. A questo fine abbiamo proposto un modello insiemistico che, come esposto nel capitolo 6, consente una migliore rappresentazione delle relazioni semantiche tra lemmi e una più facile implementabilità dell'ontologia.

Un'ontologia linguistica, per essere utilizzabile e funzionale ai suoi scopi, deve infatti potersi estendere e rappresentare le modalità specifiche con cui le lingue via via analizzate si riferiscono a questo universo. A tal fine questa tesi ha riportato una serie di casi d'uso sviluppati nel complesso lavoro di validazione di IMAGACT rispetto alla lingua cinese e li ha sfruttati come banco di prova e sorgente di requisiti concreti per la modellizzazione. I requisiti operativi e di interoperabilità dell'ontologia sono stati trattati a diversi livelli, che si legano tra loro a cascata: un modello insiemistico trova rappresentazione più semplice e facilmente implementabile in un modello di database a grafo piuttosto che in un modello relazionale e un database a grafo permette una diretta proiezione in un formato RDF, necessario per lo sfruttamento dell'ontologia all'interno di tecnologie web. Il lavoro di tesi, nel capitolo 7 ha dato conto della effettiva migrazione del DB MySQL, su cui IMAGACT è stato implementato originariamente, in un DB a grafo Neo4J e dei requisiti di adeguatezza che tale migrazione ha considerato. Nell'ultimo capitolo si è poi mostrato come il nuovo modello sviluppato in Neo4J può essere a sua volta convertito in modo semplice in un'ontologia formale e collegato ad altre risorse, quali *lemon* e BabelNet.

BIBLIOGRAFIA

A

Artstein, R. and Poesio, M. (2008). Inter-coder agreement for computational linguistics. *Comput. Linguist.*, 34(4):555–596.

B

Babko-Malaya, O., Palmer, M., Xue, N., Joshi, A., and Kulick, S. (2004). Proposition bank ii: Delving deeper. In *Frontiers in Corpus Annotation. Proceedings of the Workshop of the HLT/NAACL Conference*, pages 17–23.

Baker, C. F., Fillmore, C. J., and Lowe, J. B. (1998). The berkeley framenet project. In *Proceedings of the 17th International Conference on Computational Linguistics - Volume 1, COLING '98*, pages 86–90, Stroudsburg, PA, USA. Association for Computational Linguistics.

Bayerl, P. S. and Paul, K. I. (2011). What determines inter-coder agreement in manual annotations? a meta-analytic investigation. *Comput. Linguist.*, 37(4):699–725.

Bentivogli, L., Bocco, A., and Pianta, E. (2004). ArchiWordNet: Integrating WordNet with Domain-Specific Knowledge. pages 39–46.

Bird, S. and Liberman, M. (2000). A formal framework for linguistic annotation. *Speech Communication*, 33:23–60.

Borghi, A. M. and Cangelosi, A. (2014). Action and language integration: From humans to cognitive robots. *Topics in Cognitive Science*, 6:344—358.

Brewer, E. A. (2000). Towards robust distributed systems (abstract). In *Proceedings of the Nineteenth Annual ACM Symposium on Principles of Distributed Computing, PODC '00*, New York, NY, USA. ACM.

Brown, S. W., Rood, T., and Palmer, M. (2010). Number or nuance: Which factors restrict reliable word sense annotation? In Chair), N. C. C., Choukri, K., Maegaard, B., Mariani, J., Odijk, J., Piperidis, S., Rosner, M., and Tapias, D., editors, *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, Valletta, Malta. European Language Resources Association (ELRA).

Bruner, J. (1983). *Child's Talk: Learning to Use Language*. Norton, New York.

Busa, R. (1951). *Sancti Thomae Aquinatis hymnorum ritualium varia specimina concordantiarum: primo saggio di indici di parole automaticamente composti e stampati da macchine IBM a schede perforate*. F.lli Biccocca, Milano.

C

Cangelosi, A. and Schlesinger, M. (2015). *Developmental Robotics: From Babies to Robots*. MIT Press, Cambridge, MA.

Carletta, J. (1996). Assessing agreement on classification tasks: The kappa statistic. *Comput. Linguist.*, 22(2):249–254.

Castano, S., Ferrara, A., Lorusso, D., and Montanelli, S. (2008). On the ontology instance matching problem. In *Database and Expert Systems Application, 2008. DEXA '08. 19th International Workshop on*, pages 180–184.

Cattell, R. (2011). Scalable sql and nosql data stores. *SIGMOD Rec.*, 39(4):12–27.

Chiarcos, C., Hellmann, S., and Nordhoff, S. (2011). Towards a linguistic linked open data cloud: The open linguistics working group. *TAL*, pages 245–275.

Choi, S. and Bowerman, M. (1991). Learning to express motion events in english and korean: the influence of language-specific lexicalization patterns. *Cognition*, 41(1-3):83–121.

Cimiano, P., McCrae, J., Buitelaar, P., and Montiel-Ponsoda, E. (2013). On the role of senses in the ontology-lexicon. In Oltramari, A., Vossen, P., Qin, L., and Hovy, E., editors, *New Trends of Research in Ontologies and Lexical Resources, Theory and Applications of Natural Language Processing*, pages 43–62. Springer Berlin Heidelberg.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1):37.

Cohen, J. (1968). Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin*, 70(4):213–220.

Cresti, E. and Panunzi, A. (2013). *Introduzione ai corpora dell'italiano*. Itinerari. Linguistica. Il Mulino, Bologna.

Croft, W. and Cruse, D. (2004). *Cognitive Linguistics*. Cambridge Textbooks in Linguistics. Cambridge University Press.

D

De Felice, I., Bartolini, R., Russo, I., Quochi, V., and Monachini, M. (2014). Evaluating ImagAct-WordNet mapping for English and Italian through videos. In Basili, R., Lenci, A., and Magnini, B., editors, *Proceedings of the First Italian Conference on Computational Linguistics CLiC-it 2014 & the Fourth International Workshop EVALITA 2014*, volume I, pages 128–131. Pisa University Press.

Di Eugenio, B. and Glass, M. (2004). The kappa statistic: A second look. *Comput. Linguist.*, 30(1):95–101.

Dimitriadis, A. and Musgrave, S. (2009). Designing linguistic databases: A primer for linguists. In Everaert, M., Musgrave, S., and Dimitriadis, A., editors, *The Use of Databases in Cross-Linguistic Studies*. De Gruyter.

E

Ehrig, M. and Sure, Y. (2004). Ontology mapping – an integrated approach. In Bussler, C., Davies, J., Fensel, D., and Studer, R., editors, *The Semantic Web: Research and Applications*, volume 3053 of *Lecture Notes in Computer Science*, pages 76–91. Springer Berlin Heidelberg.

Ehrmann, M., Cecconi, F., Vannella, D., McCrae, J. P., Cimiano, P., and Navigli, R. (2014). Representing multilingual data as linked data: the case of babelnet 2.0. In *Proceedings*

of the Ninth International Conference on Language Resources and Evaluation (LREC-2014), Reykjavik, Iceland, May 26-31, 2014., pages 401–408.

F

Fellbaum, C. (1998). *WordNet: An Electronic Lexical Database*. MIT Press, Cambridge, MA.

Fillmore, C. J. and Atkins, B. T. (1992). Toward a frame-based lexicon: The semantics of risk and its neighbors. *Frames, fields, and contrasts: New essays in semantic and lexical organization*, 103.

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5):378–382.

Fox, A. and Brewer, E. (1999). Harvest, yield, and scalable tolerant systems. In *Hot Topics in Operating Systems, 1999. Proceedings of the Seventh Workshop on*, pages 174–178.

G

Gagliardi, G. (2014). *Validazione dell'ontologia dell'azione IMAGACT per lo studio e la diagnosi del Mild Cognitive Impairment*. PhD thesis, University of Florence.

Gangemi, A., Guarino, N., Masolo, C., Oltramari, A., and Schneider, L. (2002). Sweetening ontologies with dolce. In *Knowledge engineering and knowledge management: Ontologies and the semantic Web*, pages 166–181. Springer.

Gilbert, S. and Lynch, N. A. (2012). Perspectives on the cap theorem. *Computer*, 45(2):30–36.

H

Halliday, M. (1975). *Learning How to Mean*. Edward Arnold, London.

Hovy, E. and Nirenburg, S. (1992). Approximating an interlingua in a principled way. In *Proceedings of the workshop on Speech and Natural Language*, pages 261–266. Association for Computational Linguistics.

Huang, Y. and Luo, T.-j. (2014). Nosql database: A scalable, availability, high performance storage for big data. In Zu, Q., Vargas-Vera, M., and Hu, B., editors, *Pervasive Computing and the Networked World*, volume 8351 of *Lecture Notes in Computer Science*, pages 172–183. Springer International Publishing.

I

Ide, N., Fellbaum, C., Baker, C., and Passonneau, R. (2010). The manually annotated sub-corpus: A community resource for and by the people. In *Proceedings of the ACL 2010 Conference Short Papers, ACLShort '10*, pages 68–73, Stroudsburg, PA, USA. Association for Computational Linguistics.

J

Jouili, S. and Vansteenbergh, V. (2013). An empirical comparison of graph databases. In *Social Computing (SocialCom), 2013 International Conference on*, pages 708–715.

K

Kilgarriff, A. (1997). I don't believe in word senses. *Computers and the Humanities*, 31(2):91–113.

Kipper Schuler, K. (2005). *VerbNet: A broad-coverage, comprehensive verb lexicon*. PhD thesis, University of Pennsylvania.

Kipper Schuler, K., Dang, H. T., and Palmer, M. (2000). Class-based construction of a verb lexicon. In *Proceedings of the Seventeenth National Conference on Artificial Intelligence and Twelfth Conference on Innovative Applications of Artificial Intelligence*, pages 691–696. AAAI Press.

Kopecka, A. and Narasimhan, B. (2012). *Events of Putting and Taking. A crosslinguistic perspective*. John Benjamins Publishing Company, Amsterdam.

Krippendorff, K. (2004). *Content Analysis: An Introduction to Its Methodology*. Sage.

L

-
- Lenci, A., Bel, N., Busa, F., Calzolari, N., Gola, E., Monachini, M., Ogonowski, A., Peters, L., Peters, W., Ruimy, N., et al. (2000). Simple: A general framework for the development of multilingual lexicons. *International Journal of Lexicography*, 13(4):249–263.
- Levin, B. (1993). *English verb classes and alternations : a preliminary investigation*.

M

-
- Magnini, B. and Speranza, M. (2002). Merging Global and Specialized Linguistic Ontologies. In *Proceedings of the Workshop Ontolex-2002 Ontologies and Lexical Knowledge Bases, LREC-2002*, pages 43–48.
- Mark Davies, J. L. F. (1982). Measuring agreement for multinomial data. *Biometrics*, 38(4):1047–1051.
- Martin, B. T. (2007). *A Typology of Ambiguity As It Relates to Natural Language Processing*. PhD thesis, Columbia, SC, USA. AAI3296668.
- McCrae, J., Fellbaum, C., and Cimiano, P. (2014). Publishing and linking wordnet using rdf and lemon. In *Proceedings of the 3rd Workshop on Linked Data in Linguistics*.
- McCrae, J., Spohr, D., and Cimiano, P. (2011). Linking Lexical Resources and Ontologies on the Semantic Web with lemon. In *Proceedings of the 8th Extended Semantic Web Conference on The Semantic Web: Research and Applications - Volume Part I, ESWC'11*, pages 245–259, Berlin, Heidelberg. Springer-Verlag.
- Mihalcea, R. (1998). Semcor: Semantically tagged corpus. Technical report, Southern Methodist University Group of Natural Language Processing.
- Miles, A. and Brickley, D. (2005). SKOS core vocabulary specification. W3C working draft, W3C. <http://www.w3.org/TR/2005/WD-swbp-skos-core-spec-20051102>.
- Miller, G. A., Beckwith, R., Fellbaum, C., Gross, D., and Miller, K. J. (1990). Introduction to wordnet: An on-line lexical database*. *International journal of lexicography*, 3(4):235–244.
- Minsky, M. L. (1969). *Semantic Information Processing*. The MIT Press.

Bibliografia

- Moneglia, M. (1987). Senso e campi di variazione: una esplorazione sul significato di alcuni verbi causativi italiani. *Studi di Grammatica Italiana*, XIII:271–349.
- Moneglia, M. (2014a). Natural Language Ontology of Action: A Gap with Huge Consequences for Natural Language Understanding and Machine Translation. In Vetulani, Z. and Mariani, J., editors, *Human Language Technology Challenges for Computer Science and Linguistics*, volume 8387 of *Lecture Notes in Computer Science*, pages 379–395. Springer International Publishing.
- Moneglia, M. (2014b). Natural language ontology of action: A gap with huge consequences for natural language understanding and machine translation. In Vetulani, Z. and Mariani, J., editors, *Human Language Technology Challenges for Computer Science and Linguistics*, volume 8387 of *Lecture Notes in Computer Science*, pages 379–395. Springer International Publishing.
- Moneglia, M. (2014c). The semantic variation of action verbs in multilingual spontaneous speech corpora. In Raso, T. and Mello, H., editors, *Spoken Corpora and Linguistics Studies*, pages 152–190. John Benjamins Publishing Company.
- Moneglia, M., Brown, S., Frontini, F., Gagliardi, G., Khan, F., Monachini, M., and Panunzi, A. (2014a). The IMAGACT Visual Ontology. An Extendable Multilingual Infrastructure for the Representation of Lexical Encoding of Action. In Chair), N. C. C., Choukri, K., Declerck, T., Loftsson, H., Maegaard, B., Mariani, J., Moreno, A., Odijk, J., and Piperidis, S., editors, *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, Reykjavik, Iceland. European Language Resources Association (ELRA).
- Moneglia, M., Brown, S., Kar, A., Kumar, A., Ojha, A. K., Mello, H., Niharika, Jha, G. N., Ray, B., and Sharma, A. (2014b). Mapping Indian Languages onto the IMAGACT Visual Ontology of Action. In Jha, G. N., Bali, K., L, S., and Banerjee, E., editors, *Proceedings of WILDRE2 - 2nd Workshop on Indian Language Data: Resources and Evaluation at LREC'14*, Reykjavik, Iceland. European Language Resources Association (ELRA).

Bibliografia

- Moneglia, M., Frontini, F., Gagliardi, G., Russo, I., Panunzi, A., and Monachini, M. (2012a). Imagact: deriving an action ontology from spoken corpora. *Proceedings of the Eighth Joint ACL-ISO Workshop on Interoperable Semantic Annotation (isa-8)*, pages 42–47.
- Moneglia, M., Gagliardi, G., Panunzi, A., Frontini, F., Russo, I., and Monachini, M. (2012b). Imagact: Deriving an action ontology from spoken corpora. In Bunt, H., editor, *Proceedings of the Eight Joint ISO - ACL SIGSEM Workshop on Interoperable Semantic Annotation ISA-8*, pages 42–47.
- Moneglia, M., Monachini, M., Panunzi, A., Frontini, F., Gagliardi, G., and Russo, I. (2012c). Mapping a corpusinduced ontology of action verbs on italwordnet. *Proceedings of the 6th International Global WordNet Conference (GWC2012) Brno*, pages 219–226.
- Moneglia, M. and Panunzi, A. (2007). Action predicates and the ontology of action across spoken language corpora. the basic issue of the semact project. In Alcántara, M. and Declerck, T., editors, *Proceeding of the International Workshop on the Semantic Representation of Spoken Language (SRSL7)*, pages 51–58, Salamanca. Universidad de Salamanca.
- Moro, A., Cecconi, F., and Navigli, R. (2014a). Multilingual Word Sense Disambiguation and Entity Linking for Everybody. In *Proceedings of the 13th International Semantic Web Conference, Posters and Demonstrations (ISWC 2014)*, pages 25–28, Riva del Garda, Italy.
- Moro, A. and Navigli, R. (2015). SemEval-2015 Task 13: Multilingual All-Words Sense Disambiguation and Entity Linking. In *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*, pages 288–297, Denver, Colorado. Association for Computational Linguistics.
- Moro, A., Raganato, A., and Navigli, R. (2014b). Entity Linking meets Word Sense Disambiguation: a Unified Approach. *Transactions of the Association for Computational Linguistics (TACL)*, 2:231–244.

N

- Nagi, K. (2013). A new representation of wordnet using graph databases on-disk and in-memory. *International Journal on Advances in Software*, 6(3-4).

Bibliografia

- Nath, R., Seddiqui, H., and Aono, M. (2014). An efficient and scalable approach for ontology instance matching. *Journal of Computers*, 9(8).
- Navigli, R. and Ponzetto, S. P. (2012a). BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *Artificial Intelligence*, 193:217–250.
- Navigli, R. and Ponzetto, S. P. (2012b). Multilingual WSD with just a few lines of code: the BabelNet API. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (ACL 2012)*, Jeju, Korea.
- Ng, H. T., Yong, C., and Foo, K. S. (1999). A case study on Inter-Annotator agreement for word sense disambiguation. In *Proceedings of the ACL SIGLEX Workshop on Standardizing Lexical Resources (SIGLEX99)*, pages 9–13, College Park, Maryland.

O

- Otero-Cerdeira, L., Rodríguez-Martínez, F. J., and Gómez-Rodríguez, A. (2015). Ontology matching: A literature review. *Expert Systems with Applications*, 42(2):949 – 971.
- Ovchinnikova, E. (2012). *Integration of world knowledge for natural language understanding*, volume 3. Springer Science & Business Media.

P

- Palmer, M., Gildea, D., and Kingsbury, P. (2005). The proposition bank: An annotated corpus of semantic roles. *Computational linguistics*, 31(1):71–106.
- Pan, Y. (In press). *Verbi d'azione in Italiano e in Cinese Mandarino. Implementazione e validazione del cinese nell'ontologia interlinguistica dell'azione IMAGACT*. PhD thesis, Università degli Studi di Firenze.
- Panunzi, A., Felice, I. D., Gregori, L., Jacoviello, S., Monachini, M., Moneglia, M., Quochi, V., and Russo, I. (2014). Translating Action Verbs using a Dictionary of Images: the IMAGACT Ontology. In *XVI EURALEX International Congress: The User in Focus*, pages 1163–1170, Bolzano / Bozen. EURALEX 2014, EURALEX 2014.

Bibliografia

- Pease, A., Niles, I., and Li, J. (2002). The suggested upper merged ontology: A large ontology for the semantic web and its applications. In *In Working Notes of the AAAI-2002 Workshop on Ontologies and the Semantic Web*, page 2002.
- Peruzzi, A. (1982). Universali semantici: il magazzino irreperibile? *Studi di Grammatica Italiana*, XI:291–326.
- Piaget, J. (1954). *The construction of reality in the child*. Basic Books, New York.
- Pianta, E., Bentivogli, L., and Girardi, C. (2002). Multiwordnet: Developing an aligned multilingual database. In *Proceedings of the First International Conference on Global WordNet*, pages 293–302.
- Pustejovsky, J. (1991). The generative lexicon. *Computational linguistics*, 17(4):409–441.

Q

- Quine, W. (1959). Meaning and translation. In Brower, R., editor, *On Translation*. Harvard University Press.

R

- Resnik, P. and Yarowsky, D. (1999). Distinguishing systems and distinguishing senses: New evaluation methods for word sense disambiguation. *Natural language engineering*, 5(02):113–133.
- Robinson, I., Webber, J., and Eifrem, E. (2013). *Graph Databases*. O'Reilly Media, Inc.
- Roventini, A., Alonge, A., Calzolari, N., Magnini, B., and F., B. (2000). Italwordnet: a large semantic database for italian. In *Proceedings of the Second International Conference on Language Resources and Evaluation, LREC 2000, 31 May - June 2, 2000, Athens, Greece*. European Language Resources Association.
- Russo, I., Frontini, F., Felice, I. D., Khan, F., and Monachini, M. (2013). Disambiguation of Basic Action Types through Nouns ' Telic Qualia. In Saurí, R., Calzolari, N., Huang, C.-R., Lenci, A., Monachini, M., and Pustejovsky, J., editors, *Proceedings of the 6th*

Bibliografia

International Conference on Generative Approaches to the Lexicon. Generative Lexicon and Distributional Semantics, pages 70–75.

S

Sadalage, P. and Fowler, M. (2012). *NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence*. Pearson Education.

Sagri, M. T., Tiscornia, D., and Bertagna, F. (2004). Jur-WordNet. pages 305–310.

Scofield, B. (2010). Nosql – death to relational databases(?). Presentation at the CodeMash conference in Sandusky (Ohio).

Scott, W. A. (1955). Reliability of content analysis: The case of nominal scale coding. *Public Opinion Quarterly*, 19(3):321–325.

Searle, J. (1983). *Intentionality: An Essay in the Philosophy of Mind*. Cambridge paperback library. Cambridge University Press.

Studer, R., Benjamins, V. R., and Fensel, D. (1998). Knowledge engineering: Principles and methods. *Data Knowl. Eng.*, 25(1-2):161–197.

T

Tomasello, M. (2003). *Constructing a language: a usage-based theory of language acquisition*. Harvard University Press.

Tversky, A. (1977). Features of similarity. *Psychological Review*, 84(4):327–352.

V

van Dam, W., S.A., R., and Bekkering, H. (2010). How specifically are action verbs represented in the neural motor system: an fmri study. *Neuroimage*, 53(4).

Veskos, P. and Pastra, K. (2014). Perspectives of robot action generalisation: Challenges and achievements. In *Proceedings of the International IEEE Workshop on Autonomous Cognitive Robotics*, Stirling, UK.

Bibliografia

- Villegas, M. and Bel, N. (2014). Parole/simple 'lemon' ontology and lexicons. *Semantic Web journal*.
- von Wright, G. H. (1963). *Norm and Action: A Logical Enquiry*. Routledge and Kegan Paul.
- Vossen, P. (2004). Eurowordnet: a multilingual database of autonomous and language-specific wordnets connected via an inter-lingualindex. *International Journal of Lexicography*, 17(2):161–173.

W

- Warrens, M. (2010). Inequalities between multi-rater kappas. *Advances in Data Analysis and Classification*, 4(4):271–286.
- Wasow, T., Perfors, A., and Beaver, D. (2005). The puzzle of ambiguity. In Orgun, O. and Sells, P., editors, *Morphology and The Web of Grammar: Essays in Memory of Steven G. Lapointe*. CSLI Publications.
- Wittgenstein, L. (1953). *Philosophical Investigations*. Blackwell, Oxford.

RINGRAZIAMENTI

Arrivato finalmente al termine della stesura della tesi, non mi resta che ringraziare tutte le persone che mi hanno aiutato e che in qualche modo hanno contribuito a questo lavoro. Prima di tutto i miei ringraziamenti vanno al gruppo di ricerca del laboratorio LABLITA, in particolare al mio co-tutore, Prof. Massimo Moneglia, e al Prof. Alessandro Panunzi, con cui lavoro da diversi anni nell'ambito dei loro progetti di ricerca e hanno contribuito in modo sostanzioso alla realizzazione di questa tesi. Ringrazio inoltre gli studenti e i ricercatori che a vario titolo hanno avuto un ruolo attivo in questo lavoro; nello specifico desidero fare i nomi di Andrea Amelio Ravelli e Andrew Williams a cui sono grato per essersi fatti carico di annotare i dataset per la validazione delle traduzioni e di Gloria Gagliardi per i suoi lavori di validazione e, soprattutto, per aver condiviso con me il lungo lavoro di annotazione di IMAGACT.

Desidero ringraziare anche Chiara Testa per essersi presa la briga di revisionare gli errori nel testo (che non è propriamente una lettura facile e scorrevole specialmente per un chimico) e Marco Fanfani per avermi liberato dalla composizione del frontespizio in \LaTeX , che è una delle cose meno edificanti nella scrittura di una tesi.

Infine voglio dire un ultimo "grazie" ai miei genitori, Vittorio e Sonia, dal momento che non glielo dico mai e invece se lo meritano per il supporto che non mi fanno mai mancare, soprattutto in periodi stressanti come questo.