

Proceedings
of the
Second Italian Conference
on
Computational Linguistics
CLiC-it 2015

3-4 December 2015, Trento

Editors:

Cristina Bosco
Sara Tonelli
Fabio Massimo Zanzotto



aA

CLiC-it

AiC

© 2015 by AILC - Associazione Italiana di Linguistica Computazionale
sede legale: c/o Bernardo Magnini, Via delle Cave 61, 38122 Trento
codice fiscale 96101430229
email: info@ai-lc.it

Pubblicazione resa disponibile
nei termini della licenza Creative Commons
Attribuzione – Non commerciale – Non opere derivate 4.0



Accademia University Press srl
via Carlo Alberto 55
I-10123 Torino
info@aAccademia.it

isbn 978-88-99200-62-6
www.aAccademia.it/CLIC_2015

Contents

Bolzano/Bozen Corpus: Coding Information about the Speaker in IMDI Metadata Structure Marco Angster	9
Detecting the scope of negations in clinical notes Giuseppe Attardi, Vittoria Cozza, Daniele Sartiano	14
Deep Learning for Social Sensing from Tweets Giuseppe Attardi, Laura Gorrieri, Alessio Miaschi, Ruggero Petrolito.....	20
Evolution of Italian Treebank and Dependency Parsing towards Universal Dependencies Giuseppe Attardi, Simone Saletti, Maria Simi.....	25
ClT-A: un Corpus di Produzioni Scritte di Apprendenti l'Italiano L1 Annotato con Errori A. Barbagli, P. Lucisano, F. Dell'Orletta, S. Montemagni, G. Venturi	31
Deep Tweets: from Entity Linking to Sentiment Analysis Pierpaolo Basile, Valerio Basile, Malvina Nissim, Nicole Novielli.....	36
Entity Linking for Italian Tweets Pierpaolo Basile, Annalina Caputo, Giovanni Semeraro	41
Enhancing the Accuracy of Ancient Greek WordNet by Multilingual Distributional Semantics Yuri Bizzoni, Riccardo Del Gratta, Federico Boschetti, Marianne Reboul	47
Deep Neural Networks for Named Entity Recognition in Italian Daniele Bonadiman, Aliaksei Severyn, Alessandro Moschitti	51
Exploring Cross-Lingual Sense Mapping in a Multilingual Parallel Corpus Francis Bond, Giulia Bonansinga.....	56
ISACCO: a corpus for investigating spoken and written language development in Italian school-age children Dominique Brunato, Felice Dell'Orletta	62
Inconsistencies Detection in Bipolar Entailment Graphs Elena Cabrio, Serena Villata.....	67
A Graph-based Model of Contextual Information in Sentiment Analysis over Twitter Giuseppe Castellucci, Danilo Croce, Roberto Basili	72
Word Sense Discrimination: A gangplank algorithm Flavio Massimiliano Cecchini, Elisabetta Fersini.....	77
Facebook and the RealWorld: Correlations between Online and Offline Conversations Fabio Celli, Luca Polonio.....	82
La scrittura in emoji tra dizionario e traduzione Francesca Chiusaroli.....	88
On Mining Citations to Primary and Secondary Sources in Historiography Giovanni Colavizza, Frédéric Kaplan	94
Visualising Italian Language Resources: a Snapshot Riccardo Del Gratta, Francesca Frontini, Monica Monachini, Gabriella Pardelli, Irene Russo, Roberto Bartolini, Sara Goggi, Fahad Khan, Valeria Quochi, Claudia Soria, Nicoletta Calzolari.....	100

A manually-annotated Italian corpus for fine-grained sentiment analysis Marilena Di Bari, Serge Sharoff, Martin Thomas	105
From a Lexical to a Semantic Distributional Hypothesis Luigi Di Caro, Guido Boella, Alice Ruggeri, Loredana Cupi, Adebayo Kolawole, Livio Robaldo	110
An Active Learning Approach to the Classification of Non-Sentential Utterances Paolo Dragone, Pierre Lison	115
The CompWHoB Corpus: Computational Construction, Annotation and Linguistic Analysis of the White House Press Briefings Corpus Fabrizio Esposito, Pierpaolo Basile, Francesco Cutugno, Marco Venuti	120
Costituzione di un corpus giuridico parallelo italiano-arabo Fathi Fawi	125
Italian-Arabic domain terminology extraction from parallel corpora Fathi Fawi, Rodolfo Delmonte	130
Annotating opposition among verb senses: a crowdsourcing experiment Anna Feltracco, Elisabetta Jezek, Bernardo Magnini, Simone Magnolini	135
Gold standard vs. silver standard: the case of dependency parsing for Italian Michele Filannino, Marilena Di Bari	141
Phrase Structure and Ancient Anatolian languages. Methodology and challenges for a Luwian syntactic annotation Federico Giusfredi	146
Linking dei contenuti multimediali tra ontologie multilingui: i verbi di azione tra IMAGACT e BabelNet Lorenzo Gregori, Andrea Amelio Ravelli, Alessandro Panunzi	150
New wine in old wineskins: a morphology-based approach to translate medical terminology Raffaele Guarasci, Alessandro Maisto	155
Computing, memory and writing: some reflections on an early experiment in digital literary studies Giorgio Guzzetta, Federico Nanni	161
Effectiveness of Domain Adaptation Approaches for Social Media PoS Tagging Tobias Horsmann, Torsten Zesch	166
Building a Corpus on a Debate on Political Reform in Twitter Mirko Lai, Daniela Virone, Cristina Bosco, Viviana Patti	171
The OPATCH corpus platform – facing heterogeneous groups of texts and users Verena Lyding, Michel Génèreux, Katalin Szabò, Johannes Andresen	177
Generare messaggi persuasivi per una dieta salutare Alessandro Mazzei	182
FacTA: Evaluation of Event Factuality and Temporal Anchoring Anne-Lyse Minard, Manuela Speranza, Rachele Sprugnoli, Tommaso Caselli	187
TED-MWE: a bilingual parallel corpus with MWE annotation. Towards a methodology for annotating MWEs in parallel multilingual corpora Johanna Monti, Federico Sangati, Mihael Arcan	193

Digging in the Dirt: Extracting Keyphrases from Texts with KD Giovanni Moretti, Rachele Sprugnoli, Sara Tonelli.....	198
Automatic extraction of Word Combinations from corpora: evaluating methods and benchmarks Malvina Nissim, Sara Castagnoli, Francesca Masini, Gianluca E. Lebani, Lucia Passaro, Alessandro Lenci	204
Improved Written Arabic Word Parsing through Orthographic, Syntactic and Semantic constraints Nahli Ouafae, Marchi Simone.....	210
ItEM: A Vector Space Model to Bootstrap an Italian Emotive Lexicon Lucia C. Passaro, Laura Pollacci, Alessandro Lenci	215
Somewhere between Valency Frames and Synsets. Comparing Latin <i>Vallex</i> and Latin WordNet Marco Passarotti, Berta González Saavedra, Christophe Onambélé Manga.....	221
SentIta and Doxa: Italian Databases and Tools for Sentiment Analysis Purposes Serena Pelosi.....	226
Le scritture brevi dello storytelling: analisi di case studies di successo Maria Laura Pierucci	232
Tracking the Evolution of Written Language Competence: an NLP-based Approach Stefan Richter, Andrea Cimino, Felice Dell'Orletta, Giulia Venturi	236
Learning Grasping Possibilities for Artifacts: Dimensions, Weights and Distributional Semantics Irene Russo, Irene De Felice	241
Experimenting the use of catenae in Phrase-Based SMT Manuela Sanguinetti	246
Cross-language projection of multilayer semantic annotation in the NewsReaderWikinews Italian Corpus (WItaC) Manuela Speranza, Anne-Lyse Minard	252
Parsing Events: a New Perspective on Old Challenges Rachele Sprugnoli, Felice Dell'Orletta, Tommaso Caselli, Simonetta Montemagni, Cristina Bosco...	258
Generalization in Native Language Identification: Learners versus Scientists Sabrina Stehwien, Sebastian Padó.....	264
Sentiment Polarity Classification with Low-level Discourse-based Features Evgeny A. Stepanov, Giuseppe Riccardi.....	269
Analyzing and annotating for sentiment analysis the socio-political debate on #labuonascuola Marco Stranisci, Cristina Bosco, Viviana Patti, Delia Irazú Hernández Farías	274
Reference-free and Confidence-independent Binary Quality Estimation for Automatic Speech Recognition Hamed Zamani, José G. C. de Souza, Matteo Negri, Marco Turchi, Daniele Falavigna	280

Learning Grasping Possibilities for Artifacts: Dimensions, Weights and Distributional Semantics

Irene Russo*, Irene De Felice **

Istituto di linguistica Computazionale “A. Zampolli” CNR *

University of Pisa **

{irene.russo, irene.defelice}@ilc.cnr.it

Abstract

English. In this paper we want to test how grasping possibilities for concrete objects can be automatically classified. To discriminate between objects that can be manipulated with one hand and the ones that require two hands, we combine conceptual knowledge about the situational properties of the objects, which can be modeled with distributional semantic methodologies, and physical properties of the objects (i.e. their dimensions and their weights), which can be found on the web through crawling.

Italiano. *In questo articolo vogliamo testare come le possibilità di manipolazione degli oggetti concreti possano essere classificate automaticamente. Per distinguere tra oggetti che possono essere manipolati con una mano e oggetti che richiedono due mani, combiniamo conoscenza concettuale sulle proprietà situazionali dell'oggetto - rappresentandola secondo il paradigma della semantica distribuzionale - con le proprietà fisiche degli oggetti (le loro dimensioni e il loro peso) estratte dal web mediante crawling.*

1 Introduction

Distributional semantic models of word meanings are based on representations that want to be cognitively plausible and that, as a matter of fact, have been tested to produce results correlated with human judgments when concepts similarity and automatic conceptual categorizations are the aim of the experiment (Erk, 2012; Turney and Pantel, 2010).

These approaches share the idea that two nominal

concepts are similar and can be clustered in the same group if the corresponding lexemes occur in comparable linguistic contexts.

Their success is also due to the expectations of the Natural Language Processing (henceforth NLP) community: both for count and predictive models of distributional semantics (Baroni et al. 2014), the core idea is that encyclopedic knowledge packed in a big corpus can improve the performance in tasks such as word sense disambiguation.

However, purely textual representations turn out to be incomplete because in language learning and processing human beings are exposed to perceptual stimuli paired with linguistic ones: the old AI dream to ground language in the world requires the mapping between these two sources of knowledge. One of the aim of this paper is to understand how much physical knowledge can be retrieved in language. Can distributional representations of concrete nouns be helpful for the automatic classification of objects, when grasping possibilities are the focus? Could they help to discriminate between objects that can be manipulated with one hand and the ones that require two hands? More generally, how much knowledge about the physical world can be found in language?

Inspired by the cognitive psychology literature on the topic, in this paper artifactual categories are theorized as situated conceptualization where physical and situational properties meet (Barsalou 2002). These situational properties describe a physical setting or event in which the target object occurs (as *grocery store*, *fruit basket*, *slicing*, *picnic* for *apple*). In an action-based categorization of objects, these kinds of properties function as a complex relational system, which links the physical structure of the object, its use, the background settings, and the design history (Chaigneau et al. 2004). Situational properties can be derived from distributional semantic models, where each

co-occurrence vector approximates the encyclopedic knowledge about its referent.

A complementary, but more action-oriented idea, is the psychological notion of affordance as the possibilities for actions that every environmental object offers (Gibson 1979). Conceptual information concerning objects affordances can be partially acquired through language, considering verb-direct object pairs as the linguistic realizations of the relations between the actions that can be performed by an agent, and the objects involved in those actions. Affordance verbs, intended as verbs that select a distinctive action for a specific object, can be discovered through statistical measures in corpora (Russo et al. 2013).

The main assumption of this paper is that the primary affordance for grasping of an artifact largely depends on its physical properties, in particular dimensions and weight. Such features are found in e-commerce websites. Extracting these values for many similar items, for example for all instances of “plate”, may help to automatically represent average dimensions for that object. However, combining this knowledge with situational properties of objects modeled as distributional semantics vectors can help understanding if they can be combined. This issue is relevant for the implementation of a module that automatically classifies grasping possibilities for objects in embodied robotics.

The paper is structured as follow: section 2 reports on the manual annotation of grasping possibilities for a set of 143 artifacts, discussing the definition of the gold standard that will be the dataset for classification experiments in section 3. Section 4 presents conclusions and ideas for future work.

2 Manual Annotation of Grasping Possibilities

Concerning grasping possibilities for concrete objects, we expect as relevant several features. First of all, objects dimensions strongly influence the type of grasp afforded by objects. For instance, we are likely to grasp a tennis ball with a whole hand, but a soccer ball with two hands: the difference between the two spheres clearly is in their diameter.

Heavy objects require a type of grasp different from the one required by the light ones. Apart from these features, we should also consider more subjective factors, such as culture, past experience

with objects, or intentions. This is particularly evident for artifacts and tools, that are the kind of objects most typically involved in manipulation and grasping and that often have a part that is specifically designed (or more suited than others) for grasping, for its shape and conformation, such as a handle (which we may call *affording parts*; cf. De Felice, 2015; in press). However, such parts (e.g. the handle of a cup) are usually grasped when the agents intention is to use the object for its canonical function (e.g. to drink from the cup), whereas in other cases it may be ignored and a different grasp could be performed (e.g. the whole cup might be taken from the above if we simply wanted to displace it).

Therefore, we can individuate at least four different grasp types afforded by concrete entities (cf. infra): the undifferentiated one-handed or two-handed grasps; a grasp by part, i.e. directed to a specific part of the object; a grasp with instrument, for substances, aggregates or every sort of things usually manipulated with some other object.

In order to obtain a gold standard annotation of artifacts grasping possibilities, we first searched WordNet 3.0 for all the nouns that have artifact as hyperonym, obtaining a list of 1510 synsets. From this list, we chose the nouns that have enough pictures as products sold on amazon.com, since it was our intention to extract objects dimensions from this website for classification experiments (cf. 3). We selected the nouns for which at least 15 pages about that object sold on amazon.com were homogeneous - i.e. they contain objects of the same type- reducing noise caused by the crawling strategy. We obtained a total number of 143 nouns. Then, for each of these nouns, we manually annotated the type of grasp afforded by the object, according to the following classes:

- One-handed grasp: this kind of grasp is for objects that have no handles or protruding parts suited for the grasp, and that can be grasped by using only one hand. The size of two of the objects dimensions (length, width or thickness) usually does not exceed the maximum span of a hand with at least two fingers bent in order to grasp and hold something. E.g.: bowl, bottle, candle, shell, necklace, clothes peg.
- Two-handed grasp: this kind of grasp is for objects that have no handles or protruding

Table 1: Number of items per classes in the gold standard.

class	#nouns
onehand	43
onehandORpart	1
oneORtwohand	25
part	23
twohand	73
twohandORpart	3

parts suited for the grasp, and that are usually grasped with two hands, because their size exceeds the maximum span of a single hand. E.g.: board, soccer ball, player piano, table, computer.

- Grasp by part: this kind of grasp is for: (i) small or large objects that have a part specifically designed for the grasping; (ii) entities that have a well identifiable part that, even if it is not specifically designed for this specific purpose, is more suited than others for the grasping thanks to its shape and conformation. E.g. knife, jug, axe, trolley, bag.
- Grasp with instrument: this kind of grasp is mainly for substances, aggregates, and entities which cannot be (or are usually not) controlled without using some other object (an instrument, generally a container). E.g. water, broth, flour, bran, sand.

For several objects more than one grasping possibility is plausible, depending on the size (a plate can be small or big) or on the availability of a container (sand can be grasped by hand).

The dataset of 143 nouns have been annotated by two annotators and the inter-annotator agreement was 0.66. Since we need a gold standard for experiments, we managed disagreements reaching a consensus on every noun.

The gold standard contains items assigned to 6 classes, distributed as in Table 1.

3 Semantic and physical knowledge about artifacts: guessing grasping possibilities

The way humans can grasp an object can be designed as a function that depends on multiple variables, such as the presence of affording parts (i.e. handle for bag), its shape, its dimensions, its

weight and the final aim of the action of grasping, modeled here as part of the situational properties. In this paper we want to test which one of these features can help in classifying artifacts that have been manually annotated according to 6 categories (see par. 2). In particular we experiment with a combination of 4 features provided for each noun:

- distributional semantics information from two corpora (GoogleNews and instructables.com) obtained with word2vec toolkit (Mikolov et al. 2013);
- average dimensions (height, length and depth) for each object, obtained crawling at least 15 pages per object from amazon.com;
- average weight for each object, obtained crawling at least 15 pages per object from amazon.com;
- co-occurrence matrix in the corpus instructables.com with nouns that are affording parts, extracting the syntactic pattern AFFORDING PART NOUN of ARTIFACT (e.g. "handle of the bag").

Because all the big corpora available contain in general news or web crawled texts that don't mention concrete actions and concrete objects so often, we choose to build a smaller but coherent corpus of do-it-yourself instructions, with the assumption that it will contain frequent instances of concrete language.

We crawled from the website instructables.com all the titles and descriptions for the projects available online in six categories (e.g. technologies, workshop, living, food, play, outside). Cleaned of the html code, the instructables.com corpus has 17M tokens; each project was parsed with the Stanford parser (de Marneffe and Manning 2008). To test if a do-it-yourself instructions corpus is useful with respect to a generic one, we represent each noun in the following experiment as a vector extracted from GoogleNews with word2vec toolkit (Mikolov et al. 2013) but also as a vector extracted from the instructables.com corpus trained with the same toolkit. These are the purely textual representations we experimented with; to complete this knowledge we added extracted information about dimensions, weight and affording parts for 143 objects.

The list of objects' parts that afford grasping and

Table 2: Precision and recall for 8 combinations of features for the 6 classes dataset.

features	Precision	Recall
instructables.com	0.113	0.336
GoogleNews	0.113	0.336
weight	0.364	0.406
dimensions	0.413	0.517
dimensions+weight	0.561	0.531
affording parts	0.25	0.399
instructables.com + all	0.443	0.552
GoogleNews + all	0.458	0.559

are component of the pattern extracted for the feature “affording parts” has been derived with a psycholinguistic test (De Felice 2015). Thirty students of the University of Pisa were interviewed and presented with 42 images of graspable entities. For each picture, they were asked to describe in the most detailed way how they would have grasped the object represented. Among the objects depicted, there were 31 artefacts. From the interviews recorded for these artefacts, we extracted all nouns denoting objects’ parts that were named as possible target of the grasp (e.g. the handle for the bag, the cup or the ladle). The list of 78 nouns was then translated in English.

3.1 Classification Experiment

The experiment is based on a multi-label classification, since our dataset consists of 143 nouns denoting artifacts, annotated according to 6 categories. The implementation of Support Vector Multi-Classification is based on LibSVM software (Chang and Lin 2001) in WEKA with 10 fold cross-validation. Table 2 reports the results in terms of precision and recall. The best performance depends on information about average dimensions and weight of the objects. Distributional semantics vectors seems useless.

The overall performance is influenced by the fact that some classes are small in the gold standard. For this reason, we experimented with the same features including just the 91 nouns that belong to the “onehand” or “twohand” classes. In Table 3, results show again that dimensions and dimensions plus weight produce good results (with “dimensions” as the best feature), even if they do not improve the performance when combined with distributional vectors that in this case are useful per se. Again, affording parts co-occurrences pro-

Table 3: Precision and recall for 8 combinations of features on two-classes dataset (“onehand” VS “twohand”).

features	Precision	Recall
GoogleNews	0.846	0.846
weight	0.715	0.714
dimensions	0.851	0.846
dimensions+weight	0.831	0.802
affording parts	0.63	0.615
GoogleNews + all	0.846	0.846

duce the worst performance, mainly because the list of affording parts was originally derived for only 31 artefacts, and not for all the objects considered in our experiment.

4 Conclusions and Future Works

In this paper we test how distributional representations of nouns denoting artifacts can be combined with physical information about their dimensions and weights automatically extracted from an e-commerce website and with co-occurrence information about their affording parts as found in a corpus of do-it-yourself instructions. The starting hypothesis - concerning grasping possibilities as basic manipulative actions for object - was that they are conceptually a combination of situational and physical properties.

As a consequence, we expect the best performance from a mixed features models. This hypothesis is not confirmed; for the two-classes dataset (“onehand” VS “twohand”) both physical knowledge and distributional semantics vectors give good results but they don’t improve the classifier’s performance when combined.

These results are in line with the current trend to mix textual and visual features from computer vision algorithms (Bruni et al. 2012) in order to go beyond the limitations of purely textual semantic representations that cannot encode information about colors, dimensions, shapes etc. As future work we plan to integrate the features used for the experiment in this paper with representations of words as bag of visual words derived from the scale-invariant feature transform (SIFT) algorithm (Lowe 1999) that in computer vision helps to detect and describe local features in images.

References

- Barsalou, L.W. 2002. Being there conceptually: simulating categories in preparation for situated action. *Representation, Memory, and Development: Essays in Honor of Jean Mandler*, 1–15.
- Baroni, M., Dinu, G. and Kruszewski, G. 2014. Don't count, predict! A systematic comparison of context-counting vs. context-predicting semantic vectors. *Proceedings of ACL 2014 (52nd Annual Meeting of the Association for Computational Linguistics)*, East Stroudsburg PA: ACL, 238-247.
- Ashok K. Chandra, Dexter C. Kozen, and Larry J. Stockmeyer. 2012. Distributional semantics with eyes: Using image analysis to improve computational representations of word meaning. *Proceedings of ACM Multimedia*, 1219-1228.
- Chaigneau, S.E., Barsalou, L.W., and Sloman, S. 2004. Assessing the causal structure of function. *Journal of Experimental Psychology: General*, 133: 601-625.
- Chih-Chung Chang and Chih-Jen Lin. 2011. LIB-SVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1-27:27. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- Katrin Erk. 2012. Vector Space Models of Word Meaning and Phrase Meaning: A Survey. *Language and Linguistics Compass*, 6(10):635-653.
- De Felice, I. in press. Objects' parts afford action: evidence from an action description task. In V. Torrens (ed.), *Language Processing and Disorders*. Newcastle: Cambridge Scholars Publishing.
- De Felice, I. 2015. *Language and Affordances*. PhD thesis, University of Pisa, Italy.
- Gibson, J. J. 1979. *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.
- Lowe, D.G. 1999. Object recognition from local scale-invariant features. *International Conference on Computer Vision*, pp. 1150-1157.
- Marie-Catherine de Marneffe and Christopher D. Manning. 2008. The Stanford typed dependencies representation. In *COLING Workshop on Cross-framework and Cross-domain Parser Evaluation*.
- Russo, I., De Felice, I., Frontini, F., Khan, F., and Monachini, M. 2013. (Fore)seeing actions in objects. Acquiring distinctive affordances from language. In B. Sharp, and M. Zock (eds.), *Proceedings of The 10th International Workshop on Natural Language Processing and Cognitive Science - NLPCS 2013 (Marseille, France, 15-17/10/2013)*, 151-161.
- Peter D. Turney and Patrick Pantel. 2010. From Frequency to Meaning: vector space models of semantics. *J. Artif. Int. Res.*, 37(1):141-188, January.