Rough Diamonds in Natural Language Learning

David M.W. Powers and Richard Leibbrandt

AI Lab, School of Computer Science, Engineering and Mathematics
Flinders University, Bedford Park, South Australia 5042
{David.Powers,Richard.Leibbrandt}@flinders.edu.au

Abstract. Machine Learning of Natural Language provides a rich environment for exploring supervised and unsupervised learning techniques including soft clustering and rough sets. This keynote presentation will trace the course of our Natural Language Learning as well as some quite intriguing spin-off applications. The focus of the paper will be learning, by both human and computer, reinterpreting our work of the last 30 years [1-12,20-24] in terms of recent developments in Rough Sets.

Keywords: Machine Learning, Natural Language, Rough Sets, Soft Clustering, Embodied Conversational Agents, Talking Head, Thinking Head, Teaching Head, Evaluation, Informedness, Markedness, DeltaP, Information Retrieval, Visualization, Human Factors, Human Machine Interface (HxI).

1 Introduction

How does a baby learn language? What does it mean for a baby to learn language? Can we really distinguish learning language from learning about the world? Does a baby distinguish learning language from learning about the world? Does a baby learn his mother tongue? What has this to do with Rough Sets and Knowledge Technology?

1.1 Psycholinguistic Evidence about Idiolect

Research in Psycholinguistics suggests that babies do *not* learn their mother's language [3,14], and in fact formal theoretical results suggest that it is impossible for a baby to learn their mother's language in the sense of Identification in the Limit [16]. In many ways a baby's model of language or idiolect differs from that of any other person, and these differences persist into adult language. A baby's language competence resembles a rough set: there is some upper set of sentences which the baby can understand (as English say) but this is quite different from the lower set of sentences which the baby can produce. Psycholinguistic research has demonstrated that there is also an in between set of sentences the baby can imitate – generally this has been interpreted as recognition/understanding competence leading imitation/ repetition competence leading production competence. However this belies the fact that this behavior continues well beyond the classical infant language period and throughout the adult life of the individual. A rough set is arguably a better model, and the explanation is quite logical.

P. Wen et al. (Eds.): RSKT 2009, LNCS 5589, pp. 17–26, 2009.

The theory of Powers and Turk [3] is that language is more negotiated than learned, and this also bypasses the theoretical results about learnability since no specific language or grammar is being learned arbitrarily precisely. However given every one has different experiences and encounters a different set of speakers, each with their own idiolect or dialect, it is to be expected that all of these will be an influence on the learner. Moreover, the learner needs to be able to understand and communicate with a much broader range of language users than they have encountered to any given point. To some extent this is enabled by having an upper set that corresponds to something like Standard English in the region of birth and early childhood, although this will be modified dynamically as the individual moves around or others move into the dialogue space of the individual. It doesn't take any effort to understand anyone from our city of birth, or the city we have lived in for most of our lives, but the further away from our language peregrination a speaker comes, even a native speaker of our mother tongue, the more effort there is in adapting to understand and communicate. And when the conversational partner is not a native speaker and has a pronounced accent, the communication difficulties can be quite pronounced, and the time to adapt somewhat prolonged.

1.2 Early Models of Grammar and Morphology Learning

Cognitive Science has emerged as Computer Science has provided the means to model and evaluate theories from Psychology, Linguistics, Neurology and Philosophy of Mind. A pre-computational theory tends to talk about daemons and make vague claims about recognizing and competing, but the computer allows such a theory to be made precise and represented as a working computer model. Simple Computational Natural Language Learning models have been developed since the 1970s and SIGNLL and the annual CoNLL conference and its predecessor workshops [4-6,8] have been running since the start of the 1990s, providing a forum for such research.

Since the late 1960s theoretical proofs countering the negative theorems about conditions under which language cannot be learned have provided conditions under which language can be learned, including in Gold's classic paper [16] (which not only produced a negative result as to when language couldn't be learned but provided a positive loophole by which language could be learned in an unsupervised way). Models with a more statistical nature have in particular been shown [17] to be able to learn a target language arbitrarily closely – where arbitrarily closely can also be formalized in terms we would recognize as defining rough sets.

1.3 Anticipated Correction

The theory of Language Learning espoused by Powers and Turk [3] is based in large measure on Turk's idea of Anticipated Correction, which was developed as an instance that gave the lie to the Poverty of the Stimulus interpretation of the negative theoretical results about learning as meaning that language must be innate.

Anticipated Correction acknowledges that children self-correct. They start to say something, and then repair it; they finish their sentence, and then provide a corrected version, or a corrected tail of the sentence. As we speak, our recognition memory suggests that this isn't quite right – it doesn't sound the way it should, e.g. the way mother would have said it. Powers and Turk thus suggested that we maintain separate recognition and production models, the former involving a larger ‘upper set’ language that is perhaps not really formalized in terms of grammar, but more as fragments of remembered sentences and phrases, and/or their generalizations. The production of a child is what is more directly and easily available and analyzable – it is much more difficult to discover the comprehension capability of the child, as this involves active experimentation rather than mere monitoring. This better known production model is evidently grammar like, and represents a more precise ‘lower set’ language. The combination of the two is necessary to explain the psycholinguistic results as discussed earlier. But the existence of the language recognizer now allows for anticipated correction, and the potential for the lower model to be corrected when it is not, as required by theory and practice, a strict subset of the recognized language.

1.4 Linguistic Classes as Rough Sets

It is also worth noting that the idea of grammatical, morphological or phonological classes is not crisp. The exception that proves the rule is key here: we tend to have a core or lower class that is quite well defined and easily ascertainable, but at the fringes words can be pressed into service as context permits. Consider, for example, that any body part can be used as a verb for any reasonable or typical action that can be performed with it – in the absence of a preferred alternative specific to the action. Thus we can shoulder another player aside in football, we can head the ball but don’t tend to foot the ball if we mean kick – noting that there are other things we do with our foot and the ball that have other technical terms in the various codes of football, e.g. passing the ball backwards with the sole of the foot/boot. Given these specific terms are available, a child is thus likely not only to get correction from the coach or other players, but to get anticipated correction from their own recognition model.

In the case of phonological classes, consider whether /y/ is a consonant or a vowel – it is context dependent, and sometimes even when clearly pronounced as a vowel it is orthographically treated like a consonant: consider ‘played’ versus ‘playing’.

This then connects with the problem of homophony and homography – sometimes quite different meanings or sounds are reflected in the same orthography, and quite different meanings or characters are reflected by the same phonology. Generally our core lower sets may be hoped to be reasonably exclusive, but the wider upper bounds of a class may have considerable overlap of membership with other classes.

Of course it is even more complex than this. One reason for this is that the forces of language change towards efficiency modify forms that are initially distinct to become indistinct: e.g. ‘it is’ → ‘it’s’ clashing with ‘its’. Since this can happen to pronouns, which are core and defining members of a functional class of English, this negates the above desirable principle about non-overlapping cores. For example, the complex systems of h* and th* pronouns of Old and Mediaeval dialects of English were compressed into forms that lost many of the distinctions, including gender, and indeed until Caxton and his contemporaries faced the issue of standardization in the context of their publishing businesses, English usage and dialect was very inhomogeneous. This is relevant to our question of overlapping core classes even in Modern English as illustrated by the overlap of the adjectival and nominal possessive pronoun classes: {my, your, his, her, its, their} and {mine, yours, his, hers, its, theirs}.

2 Models of Grammar and Morphology Learning

2.1 Statistical and Neural Models of Language Learning

The idea of having two levels of language model also relates to a standard principle both in Machine Learning and Child Learning: *You can only learn what you almost already know*. We have to have enough hooks to hang new concepts on, and at the same time eliminate pursuing (inducing) unfruitful extrapolations that lead away from the desired language model, as represented by the upper model. This then leads further to the idea of incremental learning. For example we may be able to produce a simple NP V NP sentence (Sent→Cl) involving a transitive verb and appropriate actors and undergoers as the subject and object. But we may also recognize sentences that involve modifiers either at sentence level or at the verb level. Thus we may have in our input data and our recognition model sentences that fit Adv Cl, Cl Adv, and verb phrases (VPs) that fit Adv VP or VP Adv. The real question is though, why we don't attach the adverb (Adv) to the NP as Adv NP or NP Adv, why is it we feel they attach to either the verb (V) or else to the clause at sentence level, and that sentences are somehow verb like?

In the 1980s Powers [1,2] introduced statistics on possible ways of attaching to what we already know (as our generative grammar), including both a neural net implementation and a purely probabilistic implementation. This question also relates to the idea of V as the 'head' of the VP and thence of the clause and sentence. Note that this work doesn't assume a predefined Part-of-Speech tag (POS) and thus contrasts with the vast majority of grammar induction/learning work that is supervised at least to the extent of assuming POS is known. Rather, the child can't know a priori any POS information, and has to induce classes along with the rules. The basic idea is to decide whether to associate an unknown word left or right, and whether at the bottom level or a higher level. One or more of these may allow a hypothesis of adding the word to an existing class, whilst others may involve hypothesizing a new class. In any case for all the possible parses using hypothesized class above some threshold, statistics are collected and probabilities of classes and rules are updated.

Note that the incremental model doesn't assume it has a data set which it has to get as much as possible out of, but rather that it can adopt a take it or leave it attitude to sentences – those that are too far from what can be parsed are discarded: we limit ourselves to inducing at most one new class per level per sentence and make the ergodic assumption that future language will be similar, giving us unlimited examples.

An interesting outcome of the above statistical incremental grammar learning algorithm is that when applied to text in which the punctuation was retained as 'words', the punctuation and functional words emerged as classes, and the first complex class (rule) that emerged was essentially that sentences started with a subject (final punctuation is treated as a sentence separator and combined with a noun phrase at the start of the sentence, combining first with the article and then with the noun (or adjective then noun). This flies in the face of traditional psycholinguists, but in fact there has for a long time [18] been a variety of psycholinguistic evidence that children pay significant attention to functional words and punctuation/prosody and that this forms a framework for learning – and that from the week they are born, and probably as much as three months earlier, neonates are already recognizing key prosodic, grammatical and vocal features of their maternal language [19]. This observation motivated further work that tried to eliminate the assumptions of a basic framework that underlay the above work [1,2]: namely that we could assume a basic word structure and sentence structure. In fact, it is not clear that the concept of a word is well-defined across languages. Similarly, in psycholinguistics we tend to talk about utterances rather than sentences, since not all utterances are well formed in terms of traditional prescriptive grammars.

2.2 Rougher Classes without Segmentation Restriction

This insight about learning the simple functional classes first relates to the characteristic of the functional words and affixes (endings, inflections etc.) of a language being the main factors that distinguish languages, providing the template that allows us to fit the content words in whatever way we intend, and have the hearer understand that intent. However, we have assumed that we are given words when in fact the child hears sounds and has to decide which sounds are meant to be different and which are meant to be the same (this is the phonology of the language, with different possible sounds, called phones, being grouped together as the identically interpreted sounds or phonemes), as well as deciding which sounds group together into meaningful units of the language (this is the morphology of the language, and the units are called morphs in their individual variants and morphemes when classified together as meaning the same thing). Grouping of morphemes into words, phrases, clauses and sentences is then what we call grammar, which also includes various constraints on order and selection. The problem of grouping units appropriately is called *clustering* or *classification* (being respectively unsupervised and supervised variants on a theme), whilst *segmentation* is the problem of deciding what the units are and how they are separated out from the continuous stream of sounds.

Most natural language processing systems assume segmentation into words and sentences as a starting point, and concentrate on the structure in between these levels. This is what is normally meant by grammar or syntax. Morphology and phonology involve both classification and segmentation below the word level, whilst pragmatics and stylistics can suggest what is appropriate structure above the sentence level.

In linguistics, rewrite rules are often restricted to be binary, e.g. the transitive clause divides into Subject and Predicate, but this is essentially arbitrary – there is little in English to suggest that the subject should be divorced from the verb, rather than attached to the verb (as it is effectively in "prodrop" languages that don't require an explicit subject as it is marked by a case ending – note that even in English we can drop pronouns to telegraph "can do this; should do that"). These binary components are usually labeled with one as the "head" and the other representing a modifier, although again this nomenclature is often arbitrary. Is it the noun that requires an article, or the article that requires a noun? When we allow the functional words and affixes to specify a template for our sentences, we see that the article has the key role grammatically, but the noun still has the key role semantically. Generally the more frequent more grammaticalized elements will be combined with more contentive words that are so large a group that they are individually relatively infrequent.

An unsupervised grammar learning algorithm will make decisions about how to segment and structure into phrases, possibly guided by a branching heuristic such as this no 'more than two components' one, or even 'no more than three components'. This can also be applied below the word level to learn the aggregation of phonemes or graphemes into morphemes. Powers [4,6] generalized from the word/grammar level down to this character/morphology level and showed that sensible structure could self-organize (experimenting with both binary and ternary split restrictions). At each split the more frequent, more closed (finite) class is a functional component important to grammatical constraints and cohesion (e.g. agreement), whilst the less frequent, more open (unbounded) class is a contentive component that is more important to the semantics of the utterance. The larger composed component retains the semantic character of the more contentive component, whilst picking up the syntactic character of the functional component. It is thus not a simple head/body contrast.

The core of a class, such as vowel or noun or noun phrase, will tend to be a single unit from the previous level (/a,e,i,o,u/ are core vowels, but /ae,ai,ee,…/ are composite vowels/diphthongs; 'dog' is a core noun, and 'dogs' is a composite noun and core noun phrase, whilst 'the dogs' and the 'the big dogs' are noun phrases). This once again brings us to the idea of a core lower set of possibilities, as well as a much larger and potentially unbounded upper set of possibilities (we can add any number of adjectives, numbers, etc. in as modifiers into the noun phrase).

2.3 Discovering Frames and Classes in Child/Parent Speech

It has long been recognized that, in the speech that language-learning children hear, individual words are hugely ambiguous in their word class, or *part-of-speech*. For instance we encounter "bang" as a verb in "Don't *bang* it" and as a noun in "That was a big *bang*". Recent work in our laboratory [20-21] has focused on the problem that a child faces in finding out to which part-of-speech a particular word belongs. One obvious strategy is to take note of the context in which the word is used. Using a corpus of natural, child-directed language spoken to 18-month-olds, we identified a set of very frequent semi-abstract sentence structures in the language input, composed of function words together with slots that can accommodate many different open-class words (e.g. "Can you X that?", "There's the X", "It's a X one"). Most of these frames can be regarded as defining a noun, verb or adjective context; for instance, the frame "Can you X that?" will typically be filled by verb roots ("Can you *hear* that?", "Can you *hold* that?", "Can you *remember* that?"). These frames were clustered together to form frame classes that accepted the same sets of words into their slots, and the classes that were obtained were very similar to the "big three" traditional classes of noun, verb and adjective. Using these classes to assign parts-of-speech to words in frame context, so that words occurring in, say, a "noun" frame were classified as nouns, produced a reasonably accurate classification.

One problem with this approach is that some frames can also be ambiguous when it comes to defining a part-of-speech: the "Are you going to X?" frame can take a noun or a verb, respectively, in the utterances "Are you going to *kindy*?" vs. "Are you going to *cry*?". At the same time, there are some words that only ever occur as members of one part-of-speech; for instance, the fact that "kindy" is usually a noun could have helped with disambiguation in the case of "Are you going to kindy?". From the corpus, it is possible to identify frames and words that are associated with one part-of-speech only. These words and frames can be regarded as lower sets of "core" words and frames for that part-of-speech, while the upper sets for each part-of-speech consist of words and frames that are part-of-speech-ambiguous. The optimal way to proceed is to combine frame and word information, favoring the frame when it is a relatively unambiguous, core frame and the word when it is a core word.

Utterances were processed one by one, and words and frames received part-of-speech tags based on the information from the core sets. Every utterance was split into frame and filler words. If an unknown word occurred in a core frame, it was tagged with the part-of-speech associated with the frame, and if a core word occurred in an unknown frame, the frame was tagged with the part-of-speech of the word. Over time, some words and frames occurred with elements from more than one core set, and were tagged as being ambiguous.

It would have been problematic for words and frames to take on *all* the part-of-speech tags of words with which they co-occurred (including multiple sets of tags when the element with which they co-occurred was ambiguous). The potential pitfall is that part-of-speech ambiguity is so rife in the set of the commonest words and frames that children hear, that we would soon run the risk of lumping all words and all frames into all parts-of-speech. For this reason, the learning process obeyed a *parsimonious learning rule*. If, say, a word co-occurred in an ambiguous frame, the word took on only those parts-of-speech that were absolutely necessary, given all the frames in which that word had occurred in the past. If all of those frames, ambiguous or not, were able to accept nouns, the word remained a noun. Only if the word occurred in a frame which accepted only verbs, would the word be assigned to the verb class as well.

As processing continued, it sometimes transpired that earlier decisions about ambiguity were unwarranted. In this way, words and frames moved back and forth between the lower and upper sets of the various parts-of-speech, until convergence was reached.

During subsequent evaluation, we used the bootstrapped information to make a judgment on the part-of-speech of each filler word in each frame. In each case, a filler word was allocated to a part-of-speech if there was only one part-of-speech to which both the word and its frame could belong. The use of the parsimonious learning rule sufficed to prevent the upper sets from including all items, and in most cases, there was only one part-of-speech to which both frame and word could belong. This algorithm produced a far more accurate classification than was obtained using only the frame information.

2.4 Grounded Language Learning and Ontology Learning

No amount of learning of grammatical or morphological structure, or even of relationships between words can actually suffice to learn language, that is to communicate meaning. Children learn their understanding of the world, their world view, their social customs and cultures, in a way that is inextricably tied to their learning of language – their concurrent learning of ontology and the semantic relationships that connect mental and linguistic constructs with physical and causal relationships. Indeed a young child is not aware of which of the properties of objects, events, etc. are arbitrary (e.g. symbolic words or names for things) or inherent (e.g. intrinsic properties or parts of things). Rather our language learning systems seem to be extensions of our sensory-motor system that are designed to recognize things that hang together (associate or correlate) as either objects (whose parts stay together, whose properties and interrelationships are relatively fixed) or events (whose causes and consequences have relationships that are similarly predictable) [3]. Cognitive Linguistics is based on this idea of similarity, which in language translates to metaphor. Words which we learn in one context are extended by analogy into other contexts. Even the part-whole structure of grammar, and the constraints of word order, seem to relate to our propensity to discern part-whole structure in the world, and our expectation that they are conserved and replicated in similar ways over time.

Although working with relationships between words, with dictionaries or semantic networks, can yield performances comparable to human performance in terms of word similarity judgments [24], and our models of the strength [10,11] of such relationships concur well with results on human word association experiments, these binary connections between concepts lack the richness and depth and connection to the world that is necessary for true understanding. We have thus argued and demonstrated [1-3] that it is necessary to learn relationship using human- or robot-like sensor-motor interaction with the world, or at least some simulation of the world. This has led to us both using common representations to model language and real-world relationships, and using simulated worlds [3,12] and robot babies [9] to 'ground' our language learning and sharpen our semantic and ontological models.

3 Evaluation

Evaluation is in general very poorly done in machine learning, and in particular in natural language learning. In particular, measures are employed that do not allow for chance and bias (e.g. recall, precision and accuracy are all flawed, as are derivatives such as F-factor, as they are uninterpretable without knowing the underlying prevalence of the classes and/or the bias of the system toward particular classes). Bookmaker Informedness and Markedness [10-11] were developed to avoid these biases and have clear relationships to Correlation, Significance and the empirically derived psychological association measure DeltaP.

One of the issues with machine learning, and in particular unsupervised learning, is how to evaluate. This is especially an issue when we are dealing with fuzzy classes, rough sets or the like. On the other hand agreement on core or lower sets is likely to be higher than on an attempted self-organization of crisp clusters. Techniques developed to compare clusterings, or clusters against standards, can also be used to compare lower sets, upper sets, or lower sets with upper sets to get a measure of crispness/roughness [22]. Another kind of evaluation is human factors evaluation where we actually compare human results with those achieved by automated systems – for example, we have compared [23] keywords used to describe documents, with keywords used to search for documents, with keywords proposed as relevant by standard formulae in Information Retrieval (TFIDF variants) in a series of live human experiments.

A common mistake with machine learning is to assume that the human performance is actually better than what can be achieved automatically – we have actually exceeded human performance with our semantic similarity algorithms [24]. A related mistake is to use human judgments as gold standards – this is particularly inappropriate where the categories and rules are matters of debate and every school of linguistics has its own distinct grammar and formalism, none of which work reliably across any large corpus of text. At best we have a silver standard which we believe is mostly accurate, or a bronze standard which is a reasonable indicator but probably largely inaccurate! We therefore adopt as far as possible the strategy of evaluating performance of learned systems in real world applications with real world judgments of performance – if performance is better in (say) information retrieval or speech recognition with one grammar versus another, that provides an objective measure of its utility. For this reason, we have explored applications in Information Retrieval/Web Search [23] and Embodied Conversational Agents/Talking & Teaching Heads [12] as avenues for the exploration and evaluation of our learning technologies.

Each of these areas has different kinds of fuzziness which suits the kind of probabilistic and possibilistic models we are developing, and have indeed forced consideration of alternatives to crisp sets. The rough set is still relatively new, and the specific approach to learning it has inspired has still to be fully related to existing learning algorithms in this space. It is our hope that this paper will lead to exploration of their utility in some of the applications we have outlined.

References

1. Powers, D.M.W.: Neurolinguistics and Psycholinguistics as a Basis for Computer Acquisition of Natural Language. SIGART 84, 29–34 (1983)
2. Powers, D.M.W.: Natural Language the Natural Way. Comp. Compacts, 100–109 (1984)
3. Powers, D.M.W., Turk, C.: Machine Learning of Natural Language, Research Monograph. Springer, NewYork (1989)
4. Powers, D.M.W.: How far can self-organization go? Results in unsupervised language learning. Machine Learning of Natural Language and Ontology, 131–136 (1991)
5. Powers, D.M.W., Reeker, L. (eds.): Proceedings of the AAAI Spring Symposium on Machine Learning of Natural Language and Ontology. DFKI, Kaiserlautern (1991)
6. Powers, D.M.W.: On the Significance of Closed Classes and Boundary Conditions: Experiments in Lexical and Syntactic Learning. In: Daelemans, W., Powers, D.M.W. (eds.) Background and Experiments in Machine Learning of Natural Language: First SHOE Workshop, ITK Proceedings 92/1, Tilburg University NL, pp. 245–266 (1992)
7. Powers, D.M.W.: Unsupervised learning of linguistic structure: an empirical evaluation. Int'l Journal of Corpus Linguistics 2(11), 91–131 (1997)
8. Powers, D.M.W. (ed.): Proc. Joint Int'l Conf. on New Methods in Language Processing and Computational Natural Language Learning. ACL, Somerset (1998)
9. Powers, D.M.W.: Robot babies: what can they teach us about language acquisition? In: Leather, J., Van Dam, J. (eds.) The Ecology of Language Acquisition, pp. 160–182. Kluwer, Dordrecht (2002)
10. Powers, D.M.W.: Recall and Precision versus the Bookmaker. In: International Conference on Cognitive Science, University of New South Wales, July 2003, pp. 529–534 (2003)
11. Powers, D.W.W.: Evaluation Evaluation. In: The 18th European Conference on Artificial Intelligence (ECAI 2008), Patras, Greece, July 21-25, 2008, pp. 843–844 (2008)
12. Powers, D.M.W., Leibbrandt, R., Pfitzner, D., Luerssen, M., Lewis, T., Abrahamyan, A., Stevens, K.: Language Teaching in a Mixed Reality Games Environmen. In: Proc. 1st International Conference on PErvasive Technologies Related to Assistive Environments (PETRA). ACM International Conference Proceeding, vol. 282, Article 70, p. 7 (2008)

13. Piaget, J.: The Child's Conception of the World. Routledge and Kegan Paul, London (1928)
14. Vygotsky, L.S.: Thought and Language. MIT Press, Cambridge (1934/1962)
15. Pike, K.L.: Language in Relation to a Unified Theory of the Structure of Human Behaviour. Mouton, The Hague
16. Gold, E.M.: Language identification in the limit. Information & Control 10, 447–474 (1967)
17. Horning, J.J.: A study of grammatical inference. Proceedings of IFIP Congress 71 (1969)
18. Huey, E.B.: The psychology and pedagogy of reading. MIT Press, Cambridge (1908/1968)
19. Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., Amiel-Tison, C.: A precursor of language acquisition in young infants. Cognition 29, 143–178 (1992)
20. Leibbrandt, R.E., Powers, D.M.W.: Grammatical category induction using lexically-based templates. In: Boston Univ. Conference on Language Development, vol. 32 (2008)
21. Leibbrandt, R.E.: Part-of-speech bootstrapping using lexically-specific frames. Unpublished PhD thesis, Flinders University of South Australia (2009)
22. Pfitzner, D., Leibbrandt, R., Powers, D.M.W.: Characterization and Evaluation of Similarity Measures for Pairs of Clusterings. Knowledge and Information Systems (2008)
23. Pfitzner, D., Treharne, K., Powers, D.M.W.: User Keyword Preference: the Nwords and Rwords Experiments. Int. J. of Internet Protocol Technology 9, 149–158 (2008)
24. Yang, D., Powers, D.M.W.: Verb similarity on the taxonomy of WordNet. In: The Third International WordNet Conference (GWC 2006), pp. 121–128 (2006)