Document Categorization using Multilingual Associative Networks based on Wikipedia

Niels Bloom
Pagelink B.V.
Sherwood Rangers 29
7551 KW Hengelo
The Netherlands
n.bloom@pagelink.nl

Mariët Theune
University of Twente
PO Box 217
7500 AE Enschede
The Netherlands
m.theune@utwente.nl

Franciska De Jong
University of Twente
PO Box 217
7500 AE Enschede
The Netherlands
f.m.g.dejong@utwente.nl

ABSTRACT

Associative networks are a connectionist language model with the ability to categorize large sets of documents. In this research we combine monolingual associative networks based on Wikipedia to create a larger, multilingual associative network, using the cross-lingual connections between Wikipedia articles. We prove that such multilingual associative networks perform better than monolingual associative networks in tasks related to document categorization by comparing the results of both types of associative network on a multilingual dataset.

Categories and Subject Descriptors

I.2.6 [**Artificial Intelligence**]: Learning—*connectionism and neural nets*; I.2.7 [**Artificial Intelligence**]: Natural Language Processing—*Language parsing and understanding*

General Terms

Algorithms, Languages

Keywords

Associative Networks, Multilingual, Categorization

1. INTRODUCTION

As international corporations continue to produce ever increasing amounts of information (up to 90% of all data in the world has been generated in the past two years [11]), they build up larger and larger collections of documents about the products they sell. These documents may include technical manuals, sales brochures, and questions by clients with the answers to those questions. To structure the information and to make it easily accessible to both employees and customers directly via the web, it is helpful to classify documents into a hierarchical structure. In combination with support for concept-based browsing, this makes document collections more easily accessible, especially to users who have a generic interest (or a less clear idea) of what they need. Such users do not necessarily have a specific search query and may even just be interested in any information that is available in a general category.

A practical problem with large libraries for companies that operate in more than one country or region is that the documents may be written in multiple languages. Ideally, each document is translated into the other languages, but this option is seldom feasible. However, such multilingual collections may be useful for multilingual clients. For example, 90% of Dutch citizens speak English [1], so they would be able to use English documents, even if no Dutch translation were available. For this reason it makes sense to classify documents into a single monolingual or conceptual structure, even if the documents are in multiple languages, leaving it to the users to decide for themselves whether or not a document in a certain language is useful for their purposes.

Classifying these documents within the conceptual structure despite the different languages in which they were written poses a problem that resembles another problem encountered when working with corporate document libraries: documents within corporate libraries are often written from very different perspectives. A sales brochure will contain a very different description of the same product than a technical manual, for example. Thus, though describing the very same product, these documents will use very different terminology. This variation in terminology causes some of the more traditional methods of automatic classification to fail [7], and translation to some degree resembles this variation in terminology, as in both cases different words are used to describe the same concept.

Associative networks, described in more detail in the next section, use linguistic information such as synonymy links, part-of and is-a relations to compensate for variation in terminology [5]. Associative networks however, are only able to deal with documents in the specific language for which they are designed, as they obviously need to be programmed with the linguistic information of the specific language they cover. If we can expand associative networks to cover more than one language, their ability to compensate for different terminology might also be beneficial for dealing with variations across languages and for different terminology that stems from documents written in an entirely different language as well.

In this paper, we show that such hybrid associative networks are not only capable of handling multilingual data,

WWW 2015 Companion, May 18–22, 2015, Florence, Italy.
ACM 978-1-4503-3473-0/15/05.
http://dx.doi.org/10.1145/2740908.2743003.

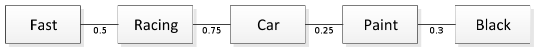

Figure 1: Simplified Associative Network

but are additionally able to produce results which are of higher quality than the results by monolingual networks.

In Section 2 we describe our method for categorizing libraries based on associative networks and in Section 3 we detail how this method can be extended to deal with multilingual document sets. Section 4 describes the experiment that we have done to compare the quality of these multilingual networks to the performance of monolingual associative networks, and we present and discuss the results in Section 5. In Section 6 we review related work. We draw our conclusions based on those results in Section 7.

2. DOCUMENT CATEGORIZATION USING ASSOCIATIVE NETWORKS

In this section we show how an associative network can be used to find connections between documents by taking one document as input, spreading activation or flow from it through the network and finding how much activation or flow each of the other documents receives. Additionally, we show how that information can be used for the purpose of document categorization.

2.1 Description of an Associative Network

An associative network is a connectionist model [3] based on work in cognitive psychology [18, 24, 25]. Like neural networks [9, 13], associative networks consist of a set of connected nodes with weights assigned to the connections. In the associative networks we use for document categorization, each node represents a concept or word, and each edge represents a conceptual connection or semantic similarity between those words. For example the word *'car'* may be associated with the words *'wheels'* and *'racing'*. Some terms may be less closely related (having a weak, low weight edge between them) while others may be more closely related (having a stronger, higher weighted edge between them). Figure 1 shows a simplified associative network consisting of five nodes with four connections, and weights for each of those connections.

When a certain word is observed in a text, other words in the network are automatically activated if they are associated with that word. This allows us to create a list of words that may be related to the text, even though they are not in the document themselves.

As an illustration, consider the sentence *'The fast black car was winning'*. Using the association model, we are able to infer information that is not explicitly provided. For example, both the word *'fast'* and the word *'car'* may be associated with *'racing'*. The word *'winning'* is also associated with *'racing'*. Thus, a document containing the sentence *'The fast black car was winning'* may be linked to the topic of *'racing'* with other documents concerning that topic, even if it does not mention the word *'racing'* at all.

This inferred information found through association is not always reliable. Association allows us to make an educated guess about information rather than providing absolute certainty; it is inferred, not deduced. However, the more words in the text we can link with a concept (such as *'racing'*), the higher the probability that this concept is adequate. We also use this fact to deal with word-sense disambiguation: when a specific surface form (such as *'fast'*) is encountered, all senses of that word are activated. Through activation of other terms in the document, more activity spreads to racing related terms and less to terms related to abstaining from food. In earlier work [5] we used various NLP techniques to eliminate incorrect word senses, but these techniques were not used in this experiment.

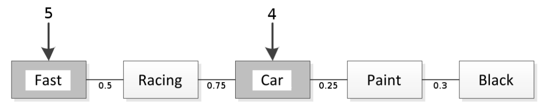

Figure 2: Input

Because the technique is applied to a very large data space, associative networks are large as well, numbering anywhere from hundreds of thousands to tens of millions of nodes. The number of connections between nodes, by comparison, is relatively small (up to a thousand edges per node), meaning the network forms a sparse graph.

Creating an associative network relies on having a source from which concepts and relations between them can be extracted for a language. One example would be WordNet, which we used in our early work [5], in which sets of synonyms can be used as nodes, and the relations between those synonyms can be used as edges. In later work [7] we described how Wikipedia can be used to create an associative network, with each article representing a node, while edges are constructed by relating the text of the articles. For more details on the way such an associative network can be constructed we refer to [7].

2.2 Spreading Activation

Once a network has been created, it can be used to make associations. To do so, it is activated by a certain input – typically a document, represented as a list of words and their frequencies in the document. For example, in Figure 2, the document used as input for the associative network contains the word *'fast'* five times and the word *'car'* four times, leading to an input value of 5 and 4 for the corresponding nodes, respectively. These input values are indicated in the figure by the incoming arrows and the marking of the nodes. The activation is spread from this input, activating neighbouring nodes in the network, which may in turn activate even more nodes.

Nodes are activated to different degrees depending on their distance to the input node and the number and weight of the connections. If the activation falls below a threshold value, the node is not activated. In Figure 3 the activation of the input spreads towards the nodes *'paint'* and *'racing'*. Since *'racing'* is connected to *'fast'* with a factor 0.5 and to *'car'* with a factor of 0.75, its activation value is $0.5 * 5 + 0.75 * 4 = 5.5$. The activation value for *'paint'* which is connected only to *'car'* by a factor of 0.25 is 1. However, the threshold value of the *'paint'* node happens to be greater than 1 (not

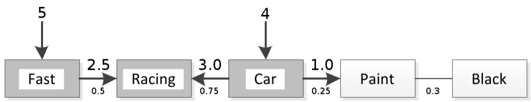

Figure 3: Activating

indicated in the figure). Thus the node is not activated (and therefore not marked in the figure) and it does not spread activation any further.

Note that in an actual network the closer two concepts are related, indicated by the weight of the edge, the more one activates the other. Thus, information will spread to closely related concepts easily while distant concepts activate one another only minimally.

Different methods can be used to calculate the exact way the activation is spread (we used Spreading Activation as described in earlier work [6]), but regardless of the method used, an activation pattern is created. This pattern is a list of the total amount of activation that was spread to each word in the associative network.

Since activation depends on both the original input and the spread, words get activated not only for being in the text, but also for being related to words in the text. Words which are directly or indirectly related to many words in the document receive spread from many sources, while concepts which are not or hardly related to the words in the document receive very little. As activation concludes a pattern is completed where concepts which are closely related to the document have a high activation while those which are not have a low activation. The pattern of a document is called its activation pattern.

Like documents, categories also have an activation pattern called a category pattern (which can be created, for example, by averaging the activation patterns of all document samples known to be in that category). By comparing the activation pattern of a document to each category pattern and calculating the distance between the two patterns by summing the difference in activation value over the terms in both patterns, we can find the category to which the document is most closely related: semantically the document belongs to the category to which it has the least distance.

3. MULTILINGUAL NETWORKS

Associative networks model relations between words that are linked by the meaning of the concept they represent, with each edge representing a specific relation. In this section we describe how this property can be used to expand an associative network to encompass more than one language.

3.1 Synonyms and Translations

Edges in the associative networks represent the conceptual connection between two nodes (word meanings) in the network. Two words could be linked because the concept represented by one word is a part of the concept represented by another such as *'car'* and *'wheel'*, because the concept represented by one word is a sub-category of the concept represented by another such as *'raven'* and *'bird'* or because the concept represented by one word is affiliated with the concept represented by another such as *'Christmas'* and *'Advent'*. However the most obvious relationship is where words are synonymous, such as *'liberty'* and *'freedom'*, which represent the same basic concept.

Based on the relationships of these synonyms, which evolved from words with similar meanings in different languages[2], we expand associative networks to cover multiple languages. For each node in the associative network, we add another node representing the translated version of the original word, connected to the original node. We can then activate either node depending on whether the original word or the translated version is present in a document, and spreading activation ensures these terms activate one another.

However, this simplistic approach presents a problem: words often do not have an absolute translation, but rather the translations are approximations that carry a slightly different meaning [20]. That meaning, sometimes described as the undertone of the word, can be quite difficult to translate [23], and can force a different impression from the original meaning which is lost in the translation. For example, Nes et al. [20] describe how the Dutch word *'wandelen'* can be translated as *'walking'*, but in turn the word *'walking'* is more accurately translated to the Dutch word *'lopen'* which is the act of walking while *'wandelen'* is generally considered to be walking for the purpose of enjoying the act. A more accurate translation of *'wandelen'* might be *'going for a walk'*, which is a bit wordier but describes the act more accurately. This shows that the word *'walk'* when translated from English to Dutch can have different meanings.

Some words do not even have an equivalent in the other language. The Dutch word *'gezellig'* describing the feeling of comfort and doing things together often in the own home might be awkwardly translated with *'cozy'*, but the latter word loses a lot of meaning compared to the original such that they really cannot be said to cover the same concept. This problem does not just exist for Dutch and English. The German word *'Schadenfreude'*, the feeling of joy or pleasure at seeing others fail or suffer misfortune, has been adopted into the English language as a loan-word as it held no common English equivalent [12]. Likewise, *'Fahrvergnügen'* was borrowed from German (which expresses the joy of driving) to be key in the 1990 U.S. Ad campaign by German car-manufacturer Volkswagen [26].

These different undertones and meanings, combined with the fact that some words are completely missing in one language or the other, mean we need a more complicated model for including a second language into an associative network.

3.2 Creating Multilingual Networks

In monolingual associative networks many of the undertones of words are already captured by links between the various words and the weights of these links. Coming back to our earlier example, the Dutch words *'wandelen'* and *'lopen'* will both be linked to *'rennen'*, as both the idea of *'going for a walk'* and of *'walking'* can be related to *'running'*. However, the link between *'wandelen'* and *'rennen'* will be less heavy than the link between *'lopen'* and *'rennen'* as the concepts of *'going for a walk'* and *'running'* are less closely related than the concepts of *'walking'* and *'running'*; the latter case is simply speeding up the motion, while the former both refer to ambulation but express different contexts in which they would be used. As monolingual networks can already capture these subtleties, one could assume that mul-

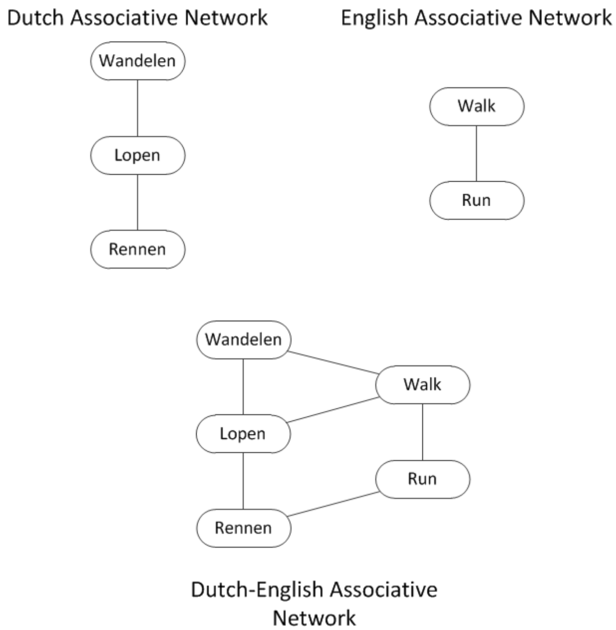

Figure 4: Combining an English and Dutch associative network

tiple monolingual networks, one for each language in the corpus could capture these differences.

Using two monolingual networks, one for each language, would certainly work when examining two independent libraries that are in different languages: each network could categorize their own language library. However, it does not allow us to group documents that cover the same topic in different languages; the two separate associative networks provide no links to one another.

To resolve the problem of grouping documents that cover the same topic in different languages, we create links between the two associative networks. This appears at first glance to bring us back to the original problem that terms cannot be translated one-to-one, but now we have a way to resolve this, as the other associative network can provide additional context to the translated term.

Moreover, instead of treating these translations as special links between two associative networks, we could treat the combined networks as a single associative network, which can be trained as a whole, rather than as two separate networks with some links between them. This also allows us to cover words with multiple possible translations – we simply link the word in one language to each of the translations, using training to set the appropriate weights between them.

In Figure 4 we display an example of two small associative networks, in this case an English and a Dutch one, being combined into a single associative network. In the first step we have two separate associative networks. In the second step, links are added between translated terms. In the final step, a new, larger associative network has been created that combines both networks, and which is ready to be trained as a whole on a multilingual dataset.

Because of the way associative networks spread activation (see Section 2), adding additional connections to nodes in the form of links between translations means the spread from one concept to a related concept becomes weaker. After all, the activation of the node now has to spread over more paths than before. Thus, it would seem at first glance that the quality of the results produced with the associative network should decrease, seeing as how less activation is spread to closely related concepts than before. However, this is not necessarily the case.

Having additional links through translated versions of words may lead to more indirect activation of a node by a related node. For example, in the original Dutch associative network in Figure 4 there is only one path from *'wandelen'* to *'lopen'*: the direct connection. However in the Dutch-English associative network, there are three paths: the direct connection, a connection through *'walk'* and a connection through *'walk'*, *'run'*, and *'rennen'*. This means that part of the activation going from *'wandelen'* to *'walk'* will loop back to *'lopen'*, which in turn means more activation spreads from *'wandelen'* to *'lopen'*.

As an alternative solution to the problem of multilingual synonyms, recent research [8, 17, 19] has been done to integrate multilingual synonyms within concept nodes. This research might be used as a basis to create and train an associative network which incorporates the subtleties of both languages directly, without the need to first create and then connect monolingual associative networks. In this research we have not examined that option, instead leaving it as future work.

4. EXPERIMENT

Multilingual associative networks allow us to categorize documents written in multiple languages. In this experiment, we examine whether a multilingual associative network produces better or worse results than monolingual associative networks.

We used Wikipedia to create our associative networks, as we did earlier in [6]. Besides providing a wealth of information and articles on a variety of topics, Wikipedia offers the additional feature that links are available between different language versions of articles, significantly simplifying the construction of a multilingual associative network.

We did not compare the associative networks to a baseline, nor use a benchmark dataset, which does not give a view of the way associative networks stand compared to state of the art methods, as we already made such comparisons in earlier work [7]. In this, we found that associative networks perform on par with Support Vector Machines while outperforming several other classification techniques [14].

4.1 Creation of the Data Sets

We created five test sets of articles within a single category. Each dataset consisted of one hundred articles, with fifty of those articles being in English and the other fifty in Dutch. The five categories were manually selected from Wikipedia, based on the criteria that they should correspond to general topics with many articles and sub-categories. The five categories selected were Animals, Biology, Chemistry, Nature and Philosophy.

Next, five sub-categories were randomly selected within each of those categories. Sub-categories which did not have at least ten articles, or which did not have a main article were excluded. Additionally, articles which had less than one thousand words, or which were marked as a stub or list article were also excluded as these would likely not contain enough information to allow for a meaningful classification. The five sub-categories were the same (or as similar as possible based on the available data) between each pair of datasets. For example, the Dutch dataset might have articles in the sub-category *'Engelse Koningen'* [1] while the English dataset would have articles in the sub-category *'English Monarchs'*.

Once these sub-categories had been selected, ten articles from each sub-category were randomly selected for a test set in each language (English and Dutch). Over five sub-categories, this thus gave two sets of 50 articles. For example, the category Biology might have Genetics, Biochemistry, Mycology, Neuroscience and Ecology as sub-categories, and thus would have ten articles from each of those sub-categories in each language. Additionally, the main article in each language was stored for each of the five sub-categories, as this was used to help classify the documents.

The articles were matched to the five possible sub-categories by comparing them with the main article for each sub-category. Dutch language articles were compared with the Dutch main article and English articles with the English main article. The category of the best matching main article was chosen as the correct sub-category.

4.2 Setup

We created three Wikipedia-based associative networks [6]: one Dutch language associative network, one English language associative network and one multilingual Dutch and English associative network combining the two monolingual ones, as described in the previous section. Each of the pair of datasets was processed by only two of the three associative networks: the associative network of the corresponding language and the multilingual associative network.

Each system was made to determine which articles matched which sub-category, based only on their textual content. No other information, such as links, was used. Similar to Bel et al. [4], an accuracy score was established based on how many articles were sorted correctly using the following formula:

$$Accuracy = \frac{CorrectlySortedArticles}{TotalArticles} * 100\% \qquad (1)$$

Thus, for example, if 40 out of 50 articles were sorted correctly, the accuracy score would be 80%.

5. RESULTS AND DISCUSSION

As can be seen in Table 1, the Dutch-English associative network performed consistently better than its monolingual equivalents. Over the 500 documents analysed, the difference in accuracy between the monolingual and multilingual associative networks is significant ($p < 0.05$).

This improvement of the multilingual associative network over the monolingual one is in line with works in cross-lingual search, such as Lavrenko et al. [15], though other work in cross-lingual text categorization such as Bel et al. [4] and Rigutini et al. [22] did not find this improvement in their work, instead observing marginally lower or equal performance for their multilingual classifiers.

This difference might be explained by the fact that the classification schemes used by Bel et al. [4] and Rigutini et al. [22] use no inherent linguistic information in their algorithms, whereas Lavrenko et al. [15] and associative networks both use information about relations between the words in the covered languages. Thus, the improvement made by using a multilingual associative network suggests that the additional connections in the associative network, as described in Section 5, help spread activation to related concepts, rather than simply smearing it out over more nodes.

It should be noted that the difference between the Dutch and the English test in terms of performance is related to the different articles as well as the different main articles for the categories in each of the two languages. For this reason, no direct comparison can be made in terms of the quality of the Dutch versus the English associative network, but the results of the monolingual and multilingual associative networks on the same dataset can be compared.

[1]Translation: English Monarchs

6. RELATED WORK

Bel et al. [4] were amongst the early pioneers examining cross-lingual text categorization. They used the Rocchio algorithm, a popular learning method based on relevance feedback, and the Winnow algorithm, a method for learning a linear classifier from labelled examples, to categorize documents in multiple languages. In earlier work [7] we compared associative networks to Rocchio amongst other classification algorithms, and found associative networks performed significantly better than Rocchio. Rigutini et al. [22] discussed extending automatic classification systems to include multiple languages, basing their work around the EM algorithm, an iterative method for finding the maximum likelihood estimates of parameters in statistical models, where the model depends on unobserved latent variables. In their work, Rigutini et al. [22] rely on translating the training data, a step that is not necessary with associative networks.

Ni et al. [21] use Wikipedia's multilingual links to extract relationships between terms in different languages for multilingual text classification. Primarily they extract relevant concepts in texts and translate these concepts through Wikipedia. We use Wikipedia's links between articles in different languages for similar purposes, effectively constructing an associative network in each language and using the relations between articles in different languages in Wikipedia to connect the two.

Our approach – combining two monolingual associative networks to create a larger bilingual one, is different from methods such as those used by Lee et al. [16], who also creates a concept-based multilingual classifier, but based on a single concept linked with word forms in multiple languages. Such methods make an assumption of a strong form of synonymy: a singular concept expressed by the corresponding words in all languages, which is not always accurate [20]. Recent research [8, 17, 19] has been done to integrate multilingual synonyms within concept nodes, and this might make such an approach for associative networks viable in the future. Our method also differs from ontology based approaches such as used by de Melo et al. [10], who use ontologies expressing more information than associative networks, which only establish how closely two concepts are related.

7. CONCLUSION AND FUTURE WORK

Multilingual associative networks provide an improvement in terms of performance over monolingual associative networks. By basing our networks on Wikipedia creating the multilingual associative network is relatively easy. Because of this, we believe it to be an effective way of increasing per-

	Dutch A.N.	Dutch-English A.N.	English A.N.
Dutch Dataset Accuracy	88%	91%	–
English Dataset Accuracy	–	86%	81%

Table 1: Results

formance of associative networks. Moreover, as Wikipedia is available in many different languages, it will be possible to make associative networks which handle many different language pairs, or even more than two languages.

In this experiment we created a multilingual network by merging two existing associative networks, but it would also be possible to create a multilingual associative network from scratch as mentioned in section 3.2, which might be an interesting experiment for future work. Trained on a multilingual data-set, this network would presumably be especially well suited for a multilingual environment, in which documents of different languages are mixed, and it may even outperform our current multilingual associative network.

8. ACKNOWLEDGEMENTS

We would like to thank Pagelink for providing the funding for this research. This publication was supported by the Dutch national program COMMIT.

9. REFERENCES

[1] Europeans and their languages. *Special Eurobarometer 386*, 2012.

[2] J. Algeo and T. Pyles. *The Origins and Development of the English Language*. Number v. 1. Wadsworth Cengage Learning, 2009.

[3] W. Bechtel. Connectionism and the philosophy of mind: an overview. *The Southern Journal of Philosophy*, 26(S1):17–41, 1988.

[4] N. Bel, C. Koster, and M. Villegas. Cross-lingual text categorization. In *Research and Advanced Technology for Digital Libraries*. Springer Berlin Heidelberg, 2003.

[5] N. Bloom. Using natural language processing to improve document categorization with associative networks. In *Proceedings of the 17th international conference on Applications of Natural Language Processing and Information Systems*, NLDB'12, pages 177–182, Berlin, Heidelberg, 2012. Springer-Verlag.

[6] N. Bloom, M. Theune, and F. de Jong. Hierarchical document categorization using associative networks. In *Proceedings of the 12th IASTED International Conference on Artificial Intelligence and Applications*, AIA 2013, February 2013.

[7] N. Bloom, M. Theune, and F. de Jong. Using wikipedia with associative networks for document classification. In *Proceedings of the 21st European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, ESANN 2013, April 2013.

[8] F. Bond and R. Foster. Linking and extending an open multilingual wordnet. In *ACL (1)*, pages 1352–1362, 2013.

[9] C. Chen, H. Lee, and C. Hwang. A hierarchical neural network document classifier with linguistic feature selection. *Applied Intelligence*, 23(3):277–294, 2005.

[10] G. de Melo and S. Siersdorfer. Multilingual text classification using ontologies. In *Advances in Information Retrieval*, pages 541–548. Springer Berlin Heidelberg, 2007.

[11] A. Dragland. Big data - for better or worse. *SINTEF Press Release*, May 2013.

[12] D. Harper. Online etymology dictionary, 2013. *http://dictionary.reference.com/browse/schadenfreude*.

[13] G. Jeschke and M. Lalmas. Hierarchical text categorisation based on neural networks and Dempster-Shafer theory of evidence. *EUROFUSE Workshop on Information Systems*, September 2002.

[14] T. Joachims. Text categorization with support vector machines: Learning with many relevant features. In *Proc. of the 10th European Conference on Machine Learning*, London, UK, 1998. Springer-Verlag.

[15] V. Lavrenko, M. Choquette, and W. B. Croft. Cross-lingual relevance models. In *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '02, pages 175–182. ACM, 2002.

[16] C. Lee and H. Yang. Construction of supervised and unsupervised learning systems for multilingual text categorization. *Expert Systems With Applications*, 36(2 Part 1):2400–2410, 2009.

[17] F. Mahdisoltani, J. Biega, and F. Suchanek. Yago3: A knowledge base from multilingual wikipedias. In *7th Biennial Conference on Innovative Data Systems Research*. CIDR 2015, 2014.

[18] G. Marcus. *The Algebraic Mind: Integrating Connectionism and Cognitive Science*. MIT Press, 2003.

[19] R. Navigli and S. P. Ponzetto. Babelnet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *Artificial Intelligence*, 193:217–250, 2012.

[20] F. Nes, T. Abma, H. Jonsson, and D. Deeg. Language differences in qualitative research: is meaning lost in translation? *European Journal of Ageing*, 7(4):313–316, 2010.

[21] C. Ni, J. Sun, J. Hu, and Z. Chen. Cross lingual text classification by mining multilingual topics from wikipedia. In *WSDM*, pages 375–384. ACM, 2011.

[22] L. Rigutini, M. Maggini, and B. Liu. In *Web Intelligence, 2005. Proceedings. The 2005 IEEE/WIC/ACM International Conference on.*

[23] L. Rohde. A case study in literary translation. *Translatio*, (2):34, 2011.

[24] R. Schank. *Dynamic Memory: A Theory of Learning in Computers and People*. Cambridge University Press, New York, 1982.

[25] R. Schank and R. Abelson. *Scripts, Plans, Goals, and Understanding*. Hillsdale, NJ: Earlbaum Assoc., 1977.

[26] R. Stanfel. Volkswagen of America, inc.: Fahrvergnügen campaign. In *Encyclopedia of Major Marketing Campaigns*. Gale Group, 2000.