Prediction of Free Word Associations Based on Hebbian Learning[1]

Reinhard Rapp and Manfred Wettler
Department of Psychology, University of Paderborn
Postfach 1621, 4790 Paderborn, Federal Republic of Germany

Abstract–An associative lexical net is built up whose weights are computed on the basis of the co-occurences of words using Hebb's rule. The co-occurences of word pairs are determined by shifting a window over a large body of text. To estimate the associative response to a given stimulus word the corresponding node is activated and its activity is propagated in the net. Our model assumes that words with high activities after propagation correspond to the associative responses of human subjects. These predictions have been tested and confirmed using the association norms collected by Russel & Jenkins from college students.

1. Introduction

A free associative response is the first word a person comes up with after listening to another word, the associative stimulus. Since Galton [4] free associations have been used to study relationships between words and what effect these relationships have on different types of behavior. Many experimental studies have shown that the associative relationships between words strongly determine the perception, learning, and forgetting of verbal material.

In this paper, we show that free associations can be explained as the result of Hebbian learning. This is done in two steps. First, the connective weights between words are computed on the basis of the frequencies and co-occurrences of these words in large bodies of texts. Then this lexical net is used to predict the associative responses of human subjects in the free association task.

The observations to be predicted are the associative responses taken from the study of Russel & Jenkins [5]. In this study 1000 subjects, American college students, were given a list of 100 stimulus words and they had to respond to each stimulus with the first word they could think of. Table 1 shows the five most frequent responses for six stimulus words.

The results of our study will help one to gain a deeper understanding of the automatic processes which determine verbal behavior; moreover, the simulation of human word associations can be an important component for natural language parsers and information retrieval systems.

	anger		baby		blue		light		ocean		working	
1	mad	353	boy	162	sky	175	dark	647	water	314	hard	132
2	fear	83	child	142	red	160	lamp	78	sea	233	loafing	99
3	hate	61	cry	113	green	125	bright	30	blue	111	sleeping	79
4	rage	49	mother	71	color	66	sun	25	deep	48	playing	65
5	temper	35	girl	51	yellow	56	bulb	23	waves	48	man	48

Table 1: Associative responses and their frequencies according to Russel & Jenkins

2. Computation of the Connective Weights

We have built up a lexical net whose nodes are the set of the 100 stimulus words of the Russel & Jenkins study, together with the five most frequent associative responses to each of the 100 stimuli. Altogether these are 371 different words.

The strengths of the connective weights have been computed on the basis of two text corpora:

1. Grolier's Electronic Encyclopedia which contains 10 million words approximately;
2. all abstracts of the psychological data-base PsycLIT, a total of 20 million words.

[1]This research was supported by the DFG (project 524/88) and by the Heinz-Nixdorf-Foundation

CH 3065-0/91/0000-0025 $1.00 © IEEE

For the calculation of the connective weights, a window is shifted from left to right over the text. After each shift the connections between all pairs of words occuring in the window are updated. Two different types of windows have been compared:

1. sentence windows which include all words in a sentence. During the process of learning they are shifted sentence by sentence;
2. continuous windows which contain a constant number of words and are shifted over the text word by word.

Seven different types of connections have been computed. The corresponding formulas are derived from different assumptions about autoassociative learning.

In formula 1 the weight between two words i and j is identical to the number of windows or sentences in which both words occur together. It is based on the assumption that the associative strength between two words increases by a constant amount whenever they appear in close neighborhood, independent of their overall frequencies. This formula gives symmetric weights with a range between zero and the number of windows or sentences.

$$w_{ij} = f \cdot P(i \,\&\, j) \tag{1}$$

where w_{ij} is the associative weight between word i and word j, f is the total number of windows, and $P(i \,\&\, j)$ is the probability of words i and j both occuring in a window.

In formula 2 the association from word i to word j is identical to the conditional probability of word j given i. Here the sum of all weights which emanate from a node tends to be the same for all the nodes of the network. This property is in agreement with the so-called fan effect, i. e. the observation of Anderson [1] and others that the amount of activation which spreads out from a node i to a node j is inversely related to the sum of all associations leading from i.

$$w_{ij} = \frac{P(i \,\&\, j)}{P(i)} \tag{2}$$

where $P(i)$ is the probability of word i occuring in a window.

Formula 3 gives the conditional probability of word i given j. Its value decreases whenever j is observed without i. This effect would be in accordance with the observation of Barnes & Underwood [2] and others that associative learning is impeded by retroactive inhibition.

$$w_{ij} = \frac{P(i \,\&\, j)}{P(j)} \tag{3}$$

Formula 4 is equal to the probability that i and j co-occur, divided by the probability of their joint co-occurrence in case of independence.

$$w_{ij} = \frac{P(i \,\&\, j)}{P(i)P(j)} \tag{4}$$

Taking the logarithm of equation 4 leads us to formula 5 defining "mutual information", which is used by Church & Hanks [3] in a similar context.

$$w_{ij} = \mathrm{ld}\frac{P(i \,\&\, j)}{P(i)P(j)} \tag{5}$$

Formulas 1 to 4 yield positive weights. Formulas 1, 4 and 5 give symmetric weights, and formulas 2 and 3 asymmetric weights with a maximum value of 1.

Formula 6 is based on the assumption that the connective weight between two words is equal to the difference between an excitatory and an inhibitory relation, the latter being reinforced whenever one word appears without the other. The derivation for this formula is given in [8].

$$w_{ij} = f \cdot \frac{P(i \,\&\, j)}{P(i)P(j)} - 1 \tag{6}$$

Formula 7 was suggested in [7] for the construction of a constraint satisfaction network that relates typical features to different types of objects.

$$w_{ij} = -\ln \frac{P(\neg i \ \& \ j)P(i \ \& \ \neg j)}{P(i \ \& \ j)P(\neg i \ \& \ \neg j)} \tag{7}$$

where $P(i \ \& \ \neg j)$ is the probability that word i but not j occurs in a window.

In formulas 1 to 7 the connective weights between two words are independent of the order of succession of these words in the text. Many studies of human memory have shown, however, that after the successive presentation of two words the association of the former to the latter word, the so-called forward association, will be stronger than its reciprocal, the so-called backward association. In order to take account of this observation we constructed, for each of the formulas 1 to 7, two variants. In the first variant, the term $P(i \ \& \ j)$ has been replaced by $P(i \rightarrow j)$, the probability that j appears in a window after i, and in the second variant $P(i \ \& \ j)$ has been replaced by $P(i \leftarrow j)$, where $P(i \ \& \ j) = P(i \rightarrow j) + P(i \leftarrow j)$.

3. Prediction of the Associative Responses

For the prediction of the associative responses of the Russel & Jenkins study, we have compared two assumptions. According to the first assumption (later on referred to as *weight model*), the associative response should be the word with the highest connective weight to the stimulus. In this case the probability that a subject responds with word j to a given stimulus i should be proportional to the connective weight w_{ij}.

According to the second assumption (the *propagation model*), the associative stimulus induces a process of propagation where an activation emanates from the stimulus word and spreads out over the lexical network in several steps. Hereby the word which receives the highest activity during propagation is considered as the associative response of the system. To test this assumption we activate, for each example separately, the node corresponding to the stimulus word and propagate this activity according to the following equation:

$$a_i(t) = input_i(t) + \sum_j a_j(t-1)w_{ij} \tag{8}$$

where $a_i(t)$ is the activity of word i after step t. $input_i(t)$is the external activation of the stimulus word and w_{ij} is the associative weight between word i and word j. w_{ii} be 0.

After each step of propagation the words of the net are ranked according to their activities. In [6] it is shown, that by the process of propagation not only first order relationships between words (direct connections) are taken into account, but also higher order relationships.

4. Results

In order to determine the validity of the different predictions we calculated, for each example separately, the ranks of the words which were given as associative responses. In the case of the weight model, the lowest rank has been given to the word with the highest connective weight to the stimulus, in the case of the propagation model to the word with the highest activity after propagation. The measure of the overall quality of a model has been defined as:

$$M = \frac{\sum_l \frac{\sum_k R_k N_k}{\sum_k N_k}}{100} \tag{9}$$

where index l denotes the 100 association examples, and k the five most frequent responses to stimulus word l in the Russel & Jenkins association experiment. N_k is the number of persons who responded with word k, R_k the rank of word k in our simulation.

M is the mean rank of the 5 most frequent human responses when looked up in the simulation word list, but weighted by the frequency of the responses and averaged over all 100 association examples. The range of M is between 1 and 371. If no correlation between predictions and observations existed, M would be 185.5, half the number of words in the network.

Formula	1	2	3	4	5	6	7
M (weight model)	49.9	49.9	44.2	44.2	44.2	46.0	44.6
M (propagation model)	49.9	49.9	35.6	37.5	40.0	38.4	39.7

Table 2: Values for M for different formulas with and without propagation. The co-occurences were determined on the basis of sentence-wise evaluation, with word order not being considered

Table 2 gives the mean ranks M when the weights were computed with different formulas on the basis of sentence-wise co-occurences in Grolier's Electronic Encyclopedia (without regard to word order). All formulas give reasonably good predictions of the associative answers, and with five of the formulas the predictions improve after propagation. It is surprising that the different methods for computing the connective weights produce similar results. The same observation has been made in earlier studies, where lexical nets have been applied to predict which query-words are used by data-bank searchers in on-line searches [8].

With an M of 35.6, formula 3 gives the best prediction of the associative responses. This is confirmed when the co-occurrences of word-pairs are determined using windows of constant width and when we take word order into account (see table 3). This result supports the assumption that the overall frequency of the response word, which is a measure of the amount of retroactive inhibition, has effect on the responses in the free association task.

	Formula	1	2	3	4	5	6	7
stimulus $\leftrightarrow$ response	M (weight model)	47.6	47.6	42.8	42.8	42.8	46.8	42.8
	M (propagation model)	47.6	47.6	35.7	36.8	37.2	37.6	36.9
stimulus $\rightarrow$ response	M (weight model)	57.9	57.9	54.2	54.2	54.2	59.8	54.5
	M (propagation model)	57.9	56.5	43.3	43.8	49.9	45.6	49.7
stimulus $\leftarrow$ response	M (weight model)	70.4	70.4	68.1	68.1	68.1	72.9	68.2
	M (propagation model)	70.4	67.8	53.9	53.2	62.3	56.2	62.2

Table 3: Values for M for different formulas with and without propagation, with order information taken into account. The co-occurences were determined using windows of width 18

When the connective weights are learned on the basis of scientific abstracts, the predictions become worse. When using sentence wise co-occurences and formula 1 we obtain a mean rank M of 69.1, which compares to an M of 49.9 when using the Grolier text corpus. This difference might be attributed to the fact that the language used in the abstracts is very specific and highly repetitive.

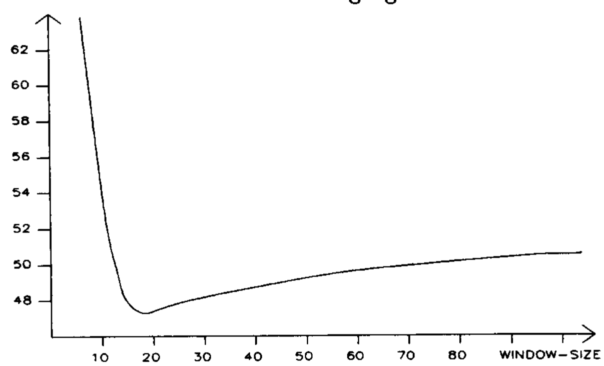

Fig. 1: M for different window-sizes

Figure 1 shows the values for M when the connective weights are computed with windows of different sizes using formula 1. The same procedure has also been applied for some of the other formulas. The optimal width of the window was always found to be around 18 words. For comparison, the average sentence length in Grolier's Electronic Encyclopedia is 22.8 words. (Non alpha-characters are counted as words.) Learning by continuously shifting windows and learning sentence by sentence give similar predictions.

Table 3 gives the values for M when the weights are learned with windows of width 18, with and without regard to word order inside the windows. The predictions are better when the direction of the connections corresponds to word order rather than vice versa. However, the best results are found when word order within the windows is not considered at all. This is probably due to the higher number of available observations when word order is neglected.

4. Conclusions

The evaluation of different sets of assumptions concerning the prediction of word associations from texts leads to the following results:

1. Type and length of the text corpus are of importance. The larger the text, i.e. the higher the absolute frequencies of the stimulus or response words, the better the predictions are. Ten million words should be considered a minimum. When we reduced our original vocabulary of 371 words to those 355 words which in Grolier's Electronic Encyclopedia have an absolute frequency of 10 or more, and removed the 16 association examples where at least one of these words occured, the mean rank M improved from 35.6 to 27.5.
2. The window size should be around 18 words, but the window type is not critical. Word order needs not be considered.
3. Different formulas for computation of the weights yield similar predictions, as long as they take the retroactive inhibition of the response word into account.
4. The predictions are improved, when second- and third-order dependencies between words are considered. For this a large vocabulary is advantageous.

When following these rules, our simulation program gives unrivalled predictions of the words people come up with in the free association task.

References

[1] Anderson, J. R.: Cognitive Psychology and its Implications. New York: Freeman & Co. 1985

[2] Barnes, J. M.; Underwood, B. J.: "Fate" of First-List Associations in Transfer Theory. Journal of Experimental Psychology, 58 (1959), 97–105

[3] Church, K. W.; Hanks, P.: Word Association Norms, Mutual Information, and Lexicography. In: Computational Linguistics, Volume 16, Number 1, March 1990

[4] Galton, F.: Psychometric Experiments. Brain 2 (1880), 149–162

[5] Jenkins, J. J.: The 1952 Minnesota Word Association Norms. In: Postman, L.; Keppel, G. (Eds.): Norms of Word Association. New York: Academic Press, 1970, 1–38.

[6] Rapp, R.; Wettler, M.: Simulation der Suchwortgenerierung im Information-Retrieval durch Propagierung in einem konnektionistischen Wortnetz. In: Nachrichten für Dokumentation 41, 1990, 27–32

[7] Rumelhart, D. E.; Smolensky, P.; McClelland, J. L.; Hinton, G. E.: Schemata and Sequential Thought Processes in PDP Models. In: Rumelhart, D. E.; McClelland, J. L. (Eds.): Parallel Distributed Processing, Vol. 2. Cambridge, MA: The MIT Press, 1986, 7–57

[8] Wettler, M.; Rapp, R.: A Connectionist System to Simulate Lexical Decisions in Information Retrieval. In: Pfeifer, R.; Schreter, Z.; Fogelman, F.; Steels, L. (Eds.): Connectionism in Perspective. Amsterdam: Elsevier Science Publishers B.V. (North-Holland) 1989.