The Effect Of Partially Relevant Information On Inferential Localisation In Decision Dichotomies

Ceyhun Aksoy
Hacettepe University

ABSTRACT

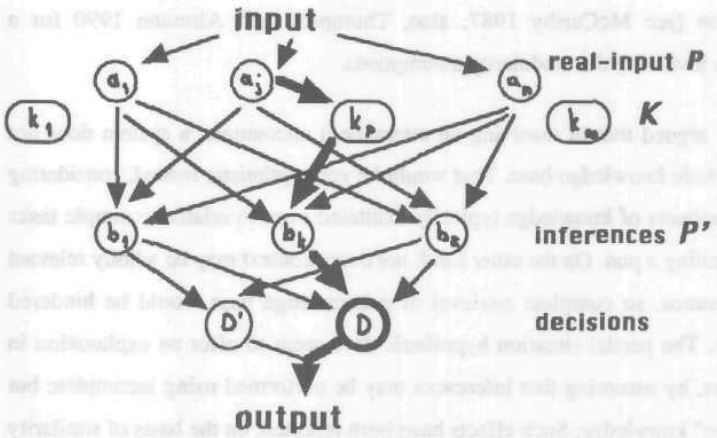
A model was proposed for the utilisation of partially-relevant information in localising an inferential unit whose eigenvector selects a decision alternative in a dichotomy. It was hypothesised that (i) a decision D related to an input text would be affected by the partialy-relevant information given previously, (ii) the decision response for D would be faster in the presence of partially-relevant information than in its absence, (iii) in the presence of partially-relevant information, the decision response for D would be faster than that for its counterpart D', (iv) the partially-relevant information and D would interact on response time. The subjects were presented a tape-recorded text with positive, negative, or no partially-relevant information prior to the input text, after which a question was asked to elicit the decision response. Response time was measured and recorded together with the response. The results supported hypotheses (i), (ii), and (iv). Implications regarding the model were discussed, and possible areas of its applicability were indicated.

Inferences are generally thought to be persisting, omni-active processes in a communication event. The inherent formalism in this approach, despite its desirable conclusions, is defeasible by the intuition that not all human communication depends on well-formed logical deductions, or is consistent (see Perlis 1987). Instead, there seem to be processes like "jumping to conclusions" effective in making default decisions (Cooper 1973). Assuming that some meaning m results across an utterance u in a context c, it is necessary to relate u to c in such a way that we can generalise m within a universe C containing all possible instances of c. Since C is an open universe by definition, the limit of c approaches infinity and that conceives an epistemic problem. The difficulty can be overcome by substituting a default quantification for c such as m in every c, which declares C as closed in m. Noticeably, the treatment yields a *partial situation* which requires further

explanation (see McCarthy 1987; also, Thompson and Altmann 1990 for a discussion involving the modularity assumption).

It may be argued that in resolving an utterance it encounters, a system does not utilise a whole knowledge base. That would be very optimistic indeed, considering the vast amounts of knowledge typically exhibited even in relatively simple tasks such as reciting a pun. On the other hand, not every context may be wholly relevant to an utterance, so complete retrieval of a knowledge base would be hindered frequently. The partial situation hypothesis does seem to offer an explanation in these terms, by assuming that inferences may be performed using incomplete but "suggestive" knowledge. Such effects have been reported, on the basis of similarity neighbourhoods (Gentner et al. 1993; Luce et al. 1990), lexical competition (Altman and Carter 1989; Marslen-Wilson 1990), lexical knowledge (Connine 1990), generic knowledge (Graesser and Clark 1985), and phonemic identification (Nakagawa 1989). *The present study is an attempt to investigate such* nonmonotonic behaviour reflected in decisions, on the basis of partial relevance.

Let us consider an input pattern p, represented at real-input level by $P\ (a_1,...,a_n)$ with vectors projecting to $P'(b_1,...,b_m)$ at inferential level. Assuming that there exists a knowledge base K which is partially relevant to p such that k_p is an associate of a_j, it can be reasoned that the activation of k_p through a_j will affect a decision based on vectors originating from a subset of P'. Such an effect may be trivial if multiple decisions are possible across P' since all elements of the inferential set will then be active by default (consult Reiter 1987). On the other hand, the effect is crucial in case of dichotomous decisions because only two elements of P' will be sustained, one of which will be imposed as output due to the activity of k_p (data-address multiplexing; see Jacobs 1988). The model can be illustrated as follows:



As can be readily noticed, k_p need not be inferentially represented because it is not a real-input item. It was assumed in this study that if a vector v passing through a_j is controlled by k_p and reaches b_k in P', then the status of a dichotomous decision D dependent on b_k will be TRUE while an alternative decision D' is assigned the status FALSE as contingent. Theoretically, this is a development of hypotheses on multistable perception (Kawamoto and Anderson 1985), assuming *gradient* processing (Golden 1986). Our assumption above can be expanded in the following predictions:

[1] b_k is *localised* by v in relation to k_p while the other elements of P' remain general;

[2] The localisation of b_k brings about the determinacy, or *fix*, of a decisional assignment in D such that

[3] D' is *masked*, understood as the replacement of the output eigenvectors of D' by those of D.

Under these assumptions, it was hypothesised that

[i] in the presence of k_p, a dichotomous decision process is resolved in favour of the localised inferential alternative, the projection of an utterance unit which is partially relevant to *K*;

[ii] response output favouring *D* in the presence of k_p is faster than response output favouring *D'* in the absence of k_p, due to the gain in vector strength obtained by the localisation process;

[iii] in the presence of k_p, response output favouring *D* is faster than that favouring *D'*, due to the masking of *D'* eigenvectors in the first case.

As response time is assumed to vary according the decision choice, an interaction between k_p and *D* was also hypothesised. The independent variable k_p was defined as a text sharing at least one predicate argument with an utterance, but whose entire propositional content cannot be represented using the given premises of the utterance. Response output was defined as the subject's choice of response as *D* or *D'*. Response output time was defined as the real time elapsed, in milliseconds, between the request for decision and the subject's response.

METHOD

Subjects. 23 female and 17 male undergraduate students, ages ranging between 20 and 24, from the Linguistics Department of Hacettepe University participated in the study. The subjects were randomly assigned to the experimental groups.

Design. A simple ANOVA design was used, employing k_p as **IV** (PPOS - present for *D*; PNEG - present for *D'*; PNIL - neutral), response time as **DV** (RT value), and decision response (D, D') as covariate.

Equipment. The experimental texts, edited using a Sound Galaxy SGBXII computer sound card operating under MS-Windows WINDAT program, were given through SONY MDR-VI headphones connected to a HI-TECH 3000E mini tape recorder, at 45 dB output. Timing controls (Tape start/stop, clock start, voice relay start) were

performed by a CHRONOS MT-1 multi-purpose time controller. Orally-given responses de-activated a CHRONOS VR-2 voice relay, which served to stop the clock, a CHRONOS TMD6 chronometer with sensitivity range 1ms.-10 mins.

Materials. Three different texts (recorded on RAKS ED-X 46 cassette tapes) were used, with contents as given in their English-translated forms below:

(1) Text1, representing *K*, according to IV categories:

(i) For PPOS,
"Faruk Yener is a rich businessman. He is famous for his kindness towards his personnel. If an employee clearly does something wrong, Faruk Yener does not announce him as guilty, but suffices with warning the person."

(ii) For PNEG,
"Faruk Yener is a rich businessman. He is famous for his cruelty towards his personnel. If an employee clearly does something wrong, Faruk Yener announces him as guilty, and immediately fires the person."

(iii) For PNIL,
"Faruk Yener is a rich businessman. He is famous for his success in overseas activities. If a foreign company applies to work with him, Faruk Yener wants 60% of the shares, and never bargains for less."

(2) Text2, representing the input,

"Cemil Bey is a clerk in a big company. The day before, he cashed a cheque for $ 100,000 on behalf of the company but the money was not registered in the treasury later on. Blamed by everyone for stealing the money, Cemil Bey was really in a very difficult position. He insisted that he would have registered the money but that he had forgotten to bring the money-sack with him. The affair was brought before the boss. The boss had just finished inspecting the matter when his very close friend, the famous businessman Faruk Yener, phoned him. When the boss finished his phone conversation, he said he had decided that Cemil Bey was *guilty /not guilty*."

In the final part of the text, the word pointing to the *guilty / not guilty* dichotomy (Turk. suç**lu** / suç**suz**, respectively) was manipulated such that the two suffixes **-lu** (affirmative) and **-suz** (negative) were superimposed by 50%-50% sampling, the

resultant pattern being rippled 30% using white noise. This was done (1) to "shell" the pragmatic context by preserving the unity of sentence structure, and (2) to prevent experimental noise by maintaining chance level for the recognition of the real-input pattern.

(3) Text3, a question to elicit the choice response:
"In your opinion, what was the boss' decision about Cemil Bey?"

Two additional texts were also used to prevent textual interference by inducing *release* activity using numeric inputs. Each release text contained 6 items from the table of random numbers (every 3rd item, left to right), given at a rate of 1 item/sec. followed by 1 sec. silence. This was seen necessary for (i) providing smooth decay in short-term memory, and (ii) preventing interference during the transfer of items into long-term memory (see Wickelgren 1973). In all texts, average playback frequency was maintained at 12kHz, based on the intuition that low sound quality would enhance attentional focusing.

Procedure. The subjects were tested individually. The materials were presented to all experimental groups in the order Text1 // Release1 // Text2 // Release2 // Text3 with time durations of 17.24, 12.0, 44.83, 12.0, and 4.43 secs., respectively. The subjects were requested to listen carefully to the texts. For release periods, the subjects were instructed to write down each number they heard minus one; e.g. to write 47, on hearing 48. The subjects' correctness in this activity was checked, and subjects with more than 3 errors per release period were discarded to ensure homoscedasticity. The onset of Text3 activated the clock and the subject's response de-activated it. The known time duration of Text3 was then subtracted from the total time elapsed and the remainder, which showed the response time, was recorded together with the decision response.

RESULTS

The data were analysed using the SPSS statistical package, MS-Windows version. The results of the analyses are given below.

The effect of partially-relevant information on decision output was found significant; the ANOVA results for this hypothesis have been summarised in the following table:

Table 1. ANOVA Results for the Effect of k_p on Decision Output

Source of Variation					
Main Effects	Sums of Squares	df	Mean Square	F	p<
k_p	3.012	2	1.506	7.974	.001
Explained	3.012	2	1.506	7.974	.001
Residual	6.988	37	.189		
Total	10.000	39	.256		

From these results, it can be concluded that k_p has a significant effect on decision response. However, the figures are unusually small because ANOVA could be run only by recoding the values of k_p as +1 (PPOS), -1 (PNEG), and 0 (PNIL), assuming equal differentiation of PPOS and PNEG from PNIL, and values of decision output as +1 (*D*) and -1 (*D'*). An analysis was also performed on the difference of means among PPOS+PNEG and PNIL on *D* (t = -3.92; df=11; p< .002), yielding a significant difference between experimental and control groups in this respect. Considering the possible existence of extreme reactions (as confirmed by the Moses Test, p< .0035), the data were re-examined by partial correlation (r = 0.5685; p< .0001), Kendall's rank-order correlation (r = 0.5073; p< .001), and Friedman's 2-way ANOVA (x^2= 34.2250; df=1; p< .0001) procedures, and finally validated using the Mann-Whitney U / Wilcoxon W (U = 35.0; W = 140.0; exact p< .0030; corrected p< .0008) and Kolmogorov-Smirnov 2-sample (Z = 1.701; p< .006) tests.

Regarding the second hypothesis predicting the effect of k_p on response time, the following results were obtained:

Table 2. ANOVA Results for the Effect of k_p on Response Time

Main Effects	Sums of Squares	Source of Variation df	Mean Square	F	p<
k_p	85617.096	2	42808.548	128.412	.0001
Explained	85617.096	2	42808.548	128.412	.0001
Residual	12334.679	37	333.370		
Total	97951.775	39	2511.584		

According to the results, k_p seems to significantly affect response time. A t-test involving PPOS+PNEG versus PNIL also supported this notion (t = -76.98; df=11; p< .0001). Further examination of the data, using partial correlation (r = - .2972; p< .033), Spearman rank-order correlation (r = -.3418; p< .031), Runs randomness (Z = -4.6452; p< .0001), Kolmogorov-Smirnov (Z = 1.701; p< .006), Mann-Whitney U / Wilcoxon W (U = 32.0, W = 269.0; p< .0024), and Friedman's 2-way ANOVA (x^2 = 40.0; df=1; p< .0001) tests, appears to support the ANOVA results.

The third research hypothesis could not be tested, due to the inadequacy of responses favouring D' in the presence of k_p (2 out of 14) for running statistical analyses. On the other hand, the testing of the hypothesis regarding the interaction between k_p and D on response time yielded the following results:

Table 3. ANOVA Results on the Interaction Between k_p and Decision Output

Main Effects	Sums of Squares	Source of Variation df	Mean Square	F	p<
k_p	85617.096	2	42808.548	164.254	.0001
D	1184.085	1	1184.085	4.543	.040
2-way interactions (k_p - D)	2289.377	2	1144.688	4.392	.020
Explained	89090.559	5	17818.112	68.367	.0001
Residual	8861.216	34	260.624		
Total	97951.775	39	2511.584		

The results seem to support the interaction hypothesis. This is also suggested by the results of non-parametric analysis employing Friedman's 2-way ANOVA (x^2= 77.11; df=2; p< .0001).

DISCUSSION

There seems to be sufficient support for three of our research hypotheses in this study. Taken up separately, the results for each hypothesis could be interpreted as follows:

[1] The significant main effect of k_p on D supports our first hypothesis, predicting that partially-relevant information affects decision responses. This is interesting in that the information that seems to prime an inference within the input domain is not propositionally related to the subdomain of a decision output by that inference. The result can be accounted for in terms of a non-orthogonality assumption, stating that vectors mapping a community in linear space need not be "far from" each other (Cooper 1973). It can also be argued that if they were, difficulties would often be encountered in recognising similar input events (consult Anderson 1983; Anderson et al. 1977). As regards our hypothesis here, it can be said that the resultant decisional activity is *induced* by the partially-relevant information. This might be an underlying process in "jumping to conclusions," where one may reason that such information becomes an *alias* argument of the input it partially relates to, but this would lead to the assumption that a deductive inference is valid in all contexts (Blakemore 1992: 14). Noticeably, more theoretical and empirical research is needed for a compromise (discussed briefly in Sperber and Wilson 1986: 179-180).

[2] The assumption that partially-relevant information is active in jumping to conclusions, supported by the non-orthogonality assumption, seems to account also for the results concerning the second hypothesis. The significant main effect of partially-relevant information on response time supports the view that the

localisation of an inferential unit enhances the output eigenvector of that unit. The eventual increment in decision speed can therefore be attributed to the presence of localisation, but the *relative* effect of partially-relevant information on localisation needs to be clarified by further research.

[3] The outcome of the hypothesis concerning the masking of alternative decision eigenvectors is left unclear. The non-testability of the hypothesis seems to be based on the strongly polarising effect of partially-relevant information on decision choices. Therefore, it seems necessary to elaborate on the *levels* of the independent variable, so that meaningful data can be obtained concerning this hypothesis.

[4] The prediction that partially-relevant information interacts with the decision response, at output level, on response time appears to be supported. Our model would expect this, due to the assumption that localisation is inference-specific and brings about a vector extention from an inference unit that is not found in other units. The decision alternative where the "extra" vector terminates, therefore, attains greater eigenvalue than its counterpart. What is critically supported by the results is that this eigenvalue is reflected in decision speed. However, the question still remains as to how the system would behave in case of optimally-triggered decisions.

CONCLUSION

In general, the present study provides support for the PDA (parallel, distributed, associative) model of verbal computations. Considering the results obtained, suspicion can be maintained that partial relevance is a *dynamic* aspect of substructural representations of utterances, and that it may be worthy of attention in studying natural language processing. Further research would make it clear how the partiality is resolved in post-decisional operations, so that a subtle distinction can be made between partially-relevant information and pre-contextual elements in intelligent communication. In this study, the former was treated as information that

can only be non-monotonically represented in terms of the utterance; i.e. virtual, in the pragmatic sense. The latter, however, would be monotonically supplemented in an utterance representation, and attain pragmatic reality.

Apart from the reality issue, questions can be raised as to the associational nature of the relations between an input pattern and a partially-relevant data source. For instance, bi-polar associations may affect a decision process in a feedforward fashion and modify the data source through feedback gain (see Usher and Zakay 1993), as in cases of biased ambiguity where the system cannot switch to an alternative solution, once having accepted a certain choice (interesting discussions by Benello et al. 1989; Kurtzman and MacDonald 1993). Or, stylistic variants of an utterance may be irretrievably modified via the selection of an expression with the highest rate of goodness-of-fit within the input set, as seems to happen in people with a tendency to use a certain variant and ignore the alternatives. The findings presented in this paper do certainly less than offer solutions, but they appear to be promising.

To conclude, it seems suitable to say that partial relevance offers a potential for future research, considering its applicability in various theoretical and practical problems in linguistics. The exploitation of the field might also be fruitful in fields of material design like CALL and AV programs, or in professional guidance literature such as expert systems.

REFERENCES

Altman, G.; D.Carter. (1989) Lexical Stress and Lexical discriminability: Stressed Syllables are More Informative, but Why? *Computer Speech and Language*. 3: 265-275.

Anderson, J.A. (1983) Cognitive and psychological computation with neural models. *IEEE Transactions on Systems*. 13(5): 799-815.

Anderson, J.A.; J.W.Silverstein; S.A. Ritz; R.S.Jones. (1977) Distinctive features, categorical perception, and probability learning: some applications of a neural model. *Psychological Review*. 84: 413-451.

Benello, J.; A.W. Mackie; J.A. Anderson. (1989) Syntactic category disambiguation with neural networks. *Computer Speech and Language*. 3:203-217.

Blakemore, D. (1992) *Understanding Utterances*. Cambridge, USA: Blackwell.

Connine, C. (1990) "Effects of sentence context and lexical knowledge in speech processing." G.T.M. Altmann (ed.) *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives*. Cambridge, Mass.: MIT Press.

Cooper, L.N. (1973) A possible organization of animal memory and learning. *Proceedings of the Nobel Symposium on Collective Properties of Physical Systems*. B.Lundquist and S.Lundquist (eds.), New York: Academic Press, pp. 252-264.

Gentner, D.; M.J. Rattermann; K.D. Forbus. (1993) The roles of similarity in transfer: Separating retrievability from inferential soundness. *Cognitive Psyhology*. 25:524-575.

Golden, R.M. (1986) The brain-state-in-a-box neural model is a gradient-descent algorithm. *Journal of Mathematical Psychology*. 30: 73-80.

Graesser, A.C.; L.F. Clark. (1985) The generation of knowledge-based inferences during narrative comprehension. G. Rickheit; H. Strohner (eds.) *Inferences in Text Processing*. Amsterdam: North-Holland.

Jacobs, R.A. (1988) Increased rates of convergence through learning rate adaptation. *Neural Networks*. 1: 295-307.

Kawamoto, A.H.; J.A. Anderson. (1985) A neural network model of multistable perception. *Acta Psygologica*. 59: 35-65.

Kurtzman, H.S.; M.C. MacDonald. (1993) Resolution of quantifier scope ambiguities. *Cognition*. 48: 243-279.

Luce, P.A.; D.B. Pisoni; S.D. Goldinger. (1990) Similarity neighborhoods of spoken words. G.T.M. Altmann (ed.) *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives*. Cambridge, Mass.: MIT Press.

Marslen-Wilson, W. (1990) Activation, competition, and frequency in lexical access. G.T.M. Altmann (ed.) *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives*. Cambridge, Mass.: MIT Press.

McCarthy, J. (1987) Epistemological problems of artificial intelligence. *Readings in Nonmonotonic Reasoning.* M.L.Ginsberg (ed.), Los Altos, Cal.: Morgan Kaufman Publishers, Inc., pp. 46-52

Nakagawa, S. (1989) Speaker-independent continuous-speech recognition by phoneme-based word spotting and time-synchronous context-free parsing. *Computer Speech and Language.* 3: 277-299.

Perlis, D. (1987) On the consistency of commonsense reasoning. *Readings in Nonmonotonic Reasoning.* M.L.Ginsberg (ed.), Los Altos, Cal.: Morgan Kaufman Publishers, Inc., pp. 56-66.

Reiter, R. (1987) A logic for default reasoning. *Readings in Nonmonotonic Reasoning.* M.L.Ginsberg (ed.), Los Altos, Cal.: Morgan Kaufman Publishers, Inc., pp. 68-92

Sperber, D.; D. Wilson (1986) *Relevance.* Oxford, UK: Blackwell.

Thompson, H.; G.T.M. Altmann. (1990) Modularity compromised: selecting partial hypotheses. G.T.M. Altmann (ed.) *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives.* Cambridge, Mass.: MIT Press.

Usher, M.; D. Zakay. (1993) A neural network model for attribute-based decision processes. *Cognitive Science.* 17: 349-396.

Wickelgren, W.A. (1973) The long and the short of memory. *Psychological Bulletin.* 80, 6: 425-437.