

# Independence of Contributing Retrieval Strategies in Data Fusion for Effective Information Retrieval

Alan F. Smeaton

School of Computer Application  
Dublin City University  
Glasnevin, Dublin 9, IRELAND  
(asmeaton@CompApp.DCU.IE)

## Abstract:

In information retrieval, data fusion is a technique for combining the outputs of more than one retrieval strategy which rank documents for retrieval. One of the observations often made about data fusion in IR is that the fusion together of document rankings from different implementations can yield a level of effectiveness which is better than either of the individual input strategies. This phenomenon has been repeatedly shown in TREC and elsewhere in IR research and it has been found in general that this holds true when the implementations are based on conceptually different approaches. In this paper we explore this hypothesis using a text retrieval application on over 200 Mbytes of Spanish newspaper texts with a fixed set of queries for which relevant documents are known. Using 9 different retrieval strategies used by different groups in TREC-4, we fuse together document rankings in different combinations in an attempt to see whether there is a correlation between the perceived conceptual independence of a document ranking strategy, and the observed improvement or otherwise in retrieval effectiveness from data fusion. Although the application we use for our experiments is text retrieval on Spanish texts, the principles we explore hold true for engineering any kind of information system based on combining the ranked retrieval of objects.

## 1. Introduction

Data fusion is a technique used for combining different sources of evidence which may be contradictory, into one unified decision for whatever the application may be. Data fusion has had much success in the area of sensors for example, in which an individual sensor may give an incorrect reading but by being combined or fused with multiple others, the result of the overall sensing operation is much more likely to be accurate and correct [1].

When used to combine more than one ranking of retrieved objects from information retrieval systems, where the rankings are based on retrieval using the same query and object representations, data fusion is a paradox. Unlike the sensor and similar applications where data fusion is used to eliminate the effects of genuinely erroneous sensor readings, or applications which fuse or combine rankings based on retrieval from different object representations (image colour, image texture, etc.), information retrieval systems fusing together the rankings from two independent retrieval algorithms operating on the same object representation can yield a fused ranking which is more effective than any of the individual input rankings. This phenomenon has been observed consistently in recent times in IR research [2-4] and the principle holds true for many kinds of information systems which are based on an inexact matching.

Information retrieval is the discipline of retrieving relevant documents from a corpus in response to a user's information need which is expressed as a query. The matching between the query and each of the documents in the corpus is an inexact match and a ranking of documents is normally presented to a user [5]. In information retrieval, data fusion can take the form of fusing the ranked output of two or more retrieval strategies or expressing the same information need as more than one query and executing the different queries using the same retrieval strategy. In this paper we explore, experimentally, the observation that data fusion works best in information retrieval when the underlying retrieval strategies being fused together are independent. We do this by fusing together in different combinations, ranked output generated using 9 different retrieval strategies on a corpus of over 200 Mbytes of text documents and attempting to correlate improvements in retrieval effectiveness with the perceived independence of the retrieval strategies being fused.

## 2. Related Work

The success of data fusion in improving retrieval performance was first observed empirically and since then there

has been work reported which attempted to rationalise and justify why this can occur. Significant work is reported by Lee [4] where he combined the outputs of 6 retrieval runs from the TREC-3 ad hoc runs and as many times before, observes how data fusion can improve retrieval performance. His analysis showed that when data fusion is shown to improve retrieval performance, the different individual retrieval runs retrieved similar sets of relevant documents but different sets of nonrelevant documents. This forms the basis for his rationale as to why data fusion works in practice.

Using a much larger set of individual document rankings (61 entries from the TREC-5 ad hoc task), Vogt [6,7] also performs pairwise combinations of retrieval rankings and analyses the resulting retrieval performances against characteristics of the individual rankings. Vogt has observed the fact that the retrieval performance of combined retrieval rankings is best when both IR systems being combined have reasonable performance figures anyway which are of similar magnitude but don't rank relevant documents in a similar fashion.

One of the common observations when data fusion in ranked retrieval proves to be effective which can be observed in the work of Lee and of Vogt and of others is that quite often the individual rankings being fused are the result of retrieval strategies which are conceptually independent of each other and thus will, by their nature, retrieve different sets of documents. This observation forms an essential part of attempts to understand and model the data fusion process as it applies to information retrieval. In this paper we set out to investigate whether the perceived conceptual independence of two retrieval strategies could be predictive of the improvement or otherwise, of combined retrieval strategies.

In the next section of the paper we describe the document collection and retrieval strategies from the TREC-4 program that we use later on and the performance of these retrieval strategies when used in isolation.

### 3. The TREC-4 Experiments on Spanish Texts

Information retrieval systems and techniques are normally compared by evaluating their effectiveness using a test collection of documents, queries and associated relevance judgements. Determining the relevance of documents where the corpus numbers tens or hundreds of thousands of documents, is beyond the reach of most IR research groups so the annual series of Text Retrieval Conferences (TREC) has grown to fill that vacuum. TREC is a worldwide coordinated benchmarking exercise for IR systems which caters for the typical IR operations of ad hoc retrieval, filtering, and many emerging specialist IR applications such as IR in languages other than English. The effect of TREC on the IR research community has been to facilitate research on large (up to 2 Gbytes or more) collections of documents and to bring the issue of scale into mainstream IR research [8].

In TREC-3, held during 1994 [9] and in TREC-4 during 1995 [10], there was a specialist track or optional line of experimentation which concentrated on information retrieval techniques applied to text documents in Spanish and using Spanish natural language queries. The document collection consisted of 208 Mbytes of full-text newspaper stories taken from the Mexican newspaper *El Norde*. There were over 68,000 individual text stories used in this, plus a set of 25 Spanish queries such as the following:

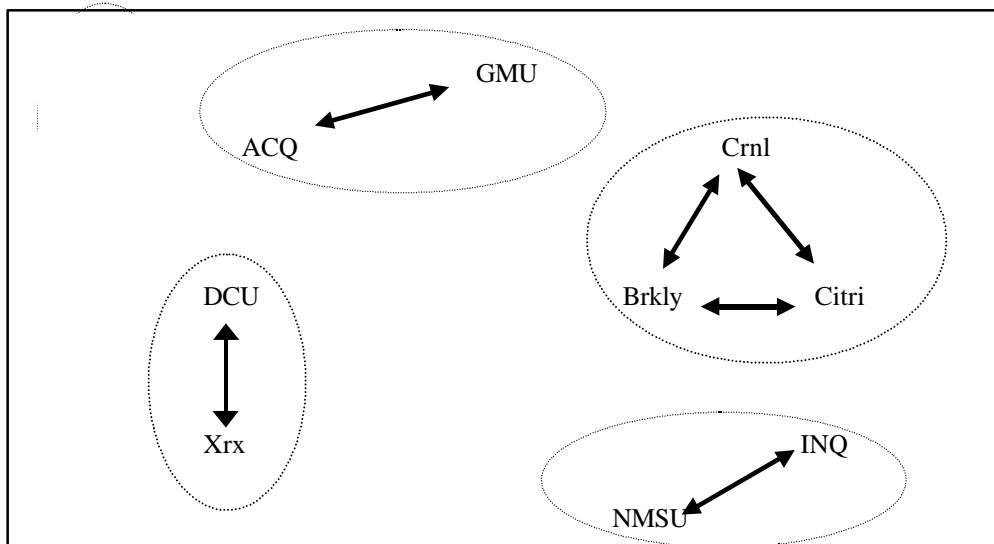
*Indicaciones de las relaciones económicas y comerciales de México con:  
los países asiáticos, por ejemplo Japon, China y Corea.*

Each group participating in Spanish TREC-4 was asked index the *El Norde* document collection using their particular approach and IR system and to run the 25 queries against their respective indexes. The top-ranked documents for each participating group per query were sent back to NIST and these were then manually assessed for relevance. When this was completed the effectiveness of the IR strategies used in retrieval for Spanish texts could be measured by calculating precision and recall figures as the relevant documents, and their rank positions in each system's ranking, were known.

In the Spanish track for TREC-4 there were 10 groups who completed the exercise. One of these, the University of Central Florida, manually hand-constructed each of the 25 queries in terms of keywords used and weights assigned to these queries. This was a labour-intensive operation requiring of the order of person-weeks of effort and that retrieval strategy, although by far the most effective one reported, would be infeasible for large-scale operational systems. Thus for this paper we restrict ourselves to working with 9 official TREC-4 runs each of which could easily be run interactively. The following is a list of those IR approaches we use in this paper including our assigned code, the name of the participating institution and a short textual description of the algorithm used. Further details of these techniques can be found in the TREC-4 Proceedings [10]

- ACQ:** Department of Defense (USA): used a vector-space model of retrieval where documents and queries were represented by n-grams of letters and the amount of n-gram overlap determined the query-document score.
- Brkly:** University of California, Berkeley (USA): included the development of a new and elaborate word stemmer for Spanish to determine indexing terms and the retrieval operation was based on a logistic regression implemented on top of the SMART vector spaced IR system.
- Citri:** CITRI/RMIT (Australia): used a simple stoplist and simple stemmer as well as word pair occurrence identification and applied these three techniques in combination with a term weighting approach.
- Crnl:** Cornell University (USA): based on the SMART implementation of the vector space model, they used term weighting and some crude but effective stemming rules and a list of stopwords.
- DCU:** Dublin City University (Ireland): applied a part-of-speech analyser to documents and queries to determine the grammatical category of each non-stopword (noun, verb, etc.) and depending on the word class, assigned greater or lesser weight to such word occurrences.
- GMU:** George Mason University (USA): implemented document retrieval on top of a parallel database machine, basing retrieval on a stoplist and a combined n-gram approach.
- NMSU:** New Mexico State University (USA): involved translating Spanish queries into English, expanding the English version of the queries with related terms and translating the expanded queries back to Spanish for execution using a version of the INQUERY system.
- INQ:** University of Massachusetts (USA): used the INQUERY system which is an implementation of a Bayesian inference network combined with a noun phrase recogniser for queries which identified phrases to be searched for in document texts, plus they also included some query expansion.
- Xrx:** Xerox Research (France and USA): used a morphological analyser, a part-of-speech tagger and a noun phrase extractor on document texts to index queries and documents by words and phrases identified from the language analysis.

Analysing these retrieval algorithms in terms of their conceptual independence of each other is difficult, a kind of black art, but there are some clear pointers we can see. The diagram in Figure 1 illustrates some of the groupings of retrieval strategies we are fusing together as we see them and thus is a first approximation of dependencies among the strategies. We can see that the Cornell and Citri approaches are similar and close to Brkly (which uses logistic regression) in that they all use similar term weighting functions within a vector space model. The DCU and Xerox approaches are similar in that they both do POS tagging of documents and queries and represent documents and queries by the baseforms of words rather than using word stems, and the ACQ and GMU approaches share the approach of using n-grams while NMSU and INQ both use Bayesian Inference nets in retrieval.



**Fig 1: Conceptual approaches of the systems fused**

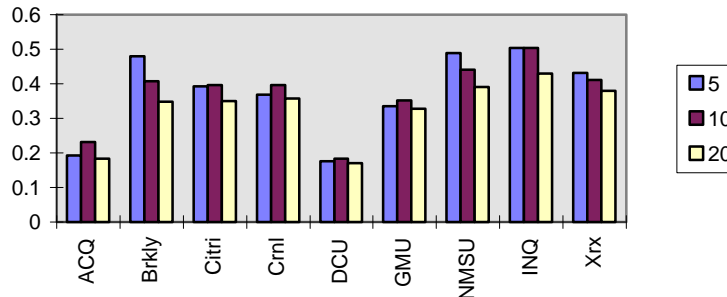
Naturally, because this is a personal grouping of strategies based on independence, it is very subjective and others may disagree with this grouping. An ideal scenario would have been to poll other researchers for their analyses and to use some grouping derived from this pooling. We do feel, however, that such a pooled grouping would not be far removed from what we have derived.

In the formal TREC experiment, algorithms and systems are evaluated based on precision-recall figures calculated on the basis of their top 1000 documents returned per query. It has often been said that the rank positions of lowly-ranked documents are meaningless in a practical context. In this work, we are interested in evaluating high-precision retrieval only, i.e. the numbers of relevant documents returned within the top rank positions. For this reason, when fusing together retrieval rankings we fuse from only the top 100 ranked documents in each ranking input. To allow fair evaluation of the effectiveness of our data fusion at high precision, instead of comparing using overall precision and recall values we use precision values at rank positions 5, 10 and 20 only and we use these 3 numbers as a basis on which to compare performances. These numbers represent the ratios of relevant documents retrieved at these rank positions, averaged over the full query set.

Table 1 shows the performances of the 9 individual baseline strategies we used for our data fusion experiments presenting figures for Precision after 5, 10 and 20 documents. This is also illustrated in Figure 1 where the 3 evaluation figures for each of the 9 retrieval strategies are plotted. These results, already published in the TREC-4 proceedings [10], show the INQ system to be the best performing individual system with the NMSU and Brkly systems just behind. Our own official runs (DCU) yielded the worst performance of this set.

	ACQ	Brkly	Citri	Crnl	DCU	GMU	NMSU	INQ	Xrx
<b>P @5</b>	0.192	0.48	0.392	0.368	0.176	0.336	0.488	0.504	0.432
<b>P @10</b>	0.232	0.408	0.396	0.396	0.184	0.352	0.44	0.504	0.412
<b>P @20</b>	0.184	0.348	0.35	0.358	0.17	0.328	0.39	0.43	0.38

**Table 1: Performance Figures for Baseline Systems**



**Figure 2: Performance Figures for Baseline Systems showing Precision at Different Rank Positions**

The next section will describe our data fusion experiments but it is important to re-emphasise that the work described in this paper is about data fusion and how the independence of the retrieval strategies affects retrieval performance; it is not about the effectiveness of information retrieval on a collection of Spanish text documents as the information retrieval application is simply the vehicle for the investigation into data fusion. Equally, other information retrieval systems and data sets could have been used. What makes information retrieval and in particular the Spanish data collection attractive for this work is that thorough relevance judgments have been made on the top 100 documents returned for each system ranking. Furthermore, our fusion techniques will only retrieve documents whose relevance to the query in question will have been manually judged as part of TREC anyway thus the evaluation of our data fusion performance will be faithful and true. Finally, and importantly for this work, with the TREC-4 Spanish results we have 9 retrieval algorithms which are very heterogeneous and which represent a good spread of IR techniques and this allows us to explore the conditions under which data fusion seems to work

best.

The purpose of the experiments reported here is to explore this relationship between the independence of the underlying retrieval strategies generating document rankings, and the improvement or otherwise of the fusion of those rankings. In the next section, we describe the experiments we have carried out to do this.

### 4. Experimental Results from Data Fusion

The data fusion paradox to emerge in information retrieval research is that the fusion together of two or more independent retrieval strategies often leads to a combined level of effectiveness which is greater than the effectiveness obtained when the strategies are used in isolation. This means that if we take a query and run it on a collection of documents using retrieval strategy A to generate a ranking of documents and then run the same query on the same collection of documents to generate a different ranking using strategy B and then we combine or fuse together the two rankings by adding normalised document scores, the effectiveness of the fused ranking will generally be better than the ranking generated by either A or B, if A and B are independent. While this observation does not always hold true there is convincing evidence in information retrieval research that this phenomenon occurs with reasonable regularity.

In most IR systems a ranking of documents is normally achieved based on computing a score for each document in the collection as an estimate of its likely relevance to the query. For data fusion of such document rankings, either the absolute rank positions of a given document in two or more rankings, or the normalised scores of a document according to two or more retrieval strategies, can be combined, typically by summation. In our work described here we fuse together rankings based on normalised document scores, where the scores assigned to all documents in all system rankings for all queries are mapped into the range [0..1] with the value 1 assigned to the highest-scored document. In some experiments reported in [4], Lee has compared the effectiveness of data fusion when combining using document rank against fusion using query-document similarity and has found that using similarity value provides slightly better retrieval effectiveness but the difference is very minor (1.3%). We choose to combine rankings based on document scores for the simple reason that computing document rank positions at the time of data fusion requires document scores to be sorted in order to generate the rank positions. It may be possible to compute and combine document scores without having to sort RSVs and this can save considerable processing costs.

Our first experiment was a pair-wise fusion of all combinations of the 9 retrieval strategies and these results are shown in Figures 2 to 10 in an appendix where the actual precision values we obtained are also given. Each of the charts in Figures 2 to 10 show the Precision values at rank positions 5, 10 and 20 for the fusion between a given retrieval strategy and each of the other 8. The charts also include the performance of the given retrieval strategy used in isolation (i.e. fused with itself).

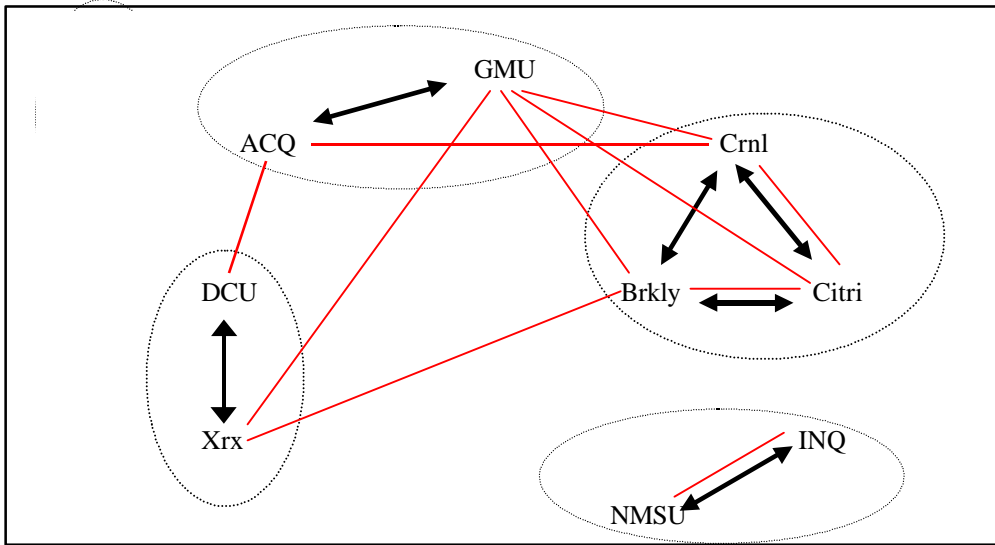
It is important to re-emphasise here that what we are doing is combining the outputs of retrieval rankings in order to see if we observe improvements in retrieval effectiveness above **both** of the input rankings and in Table 2 we highlight such cases of improvements with shaded boxes.

	ACQ	Brkly	Citri	Crnl	DCU	GMU	NMSU	INQ	Xrx
ACQ									
Brkly									
Citri									
Crnl									
DCU									
GMU									
NMSU									
INQ									
Xrx									

**Table 2: Cases where a pair-wise fusion is more effective than both of its inputs**

Table 2 shows there are only 10 cases of improvement in retrieval effectiveness out of a possible 36 symmetric pairwise combinations. Among these 10 there are some groupings as follows: As expected, because it is so different to the other strategies, GMU seems to be the most independent of each of the others, causing improved effectiveness when fused with 4 others (Brkly, Citri, Crnl and Xrx) and improving 2 more over their own individual baselines (ACQ and DCU). Unexpectedly, neither INQ nor NMSU fuse well with any others except with each other and when they are fused we get the best level of performance of any run in the series. This is initially a surprise since as we pointed out earlier, these are the only systems to use inference nets. However this may be explained by the fact that both approaches use different kinds of query expansion and thus the queries used may be very different. The ACQ and DCU yield improvement over baselines only when fused with each other and when ACQ is fused with Crnl, possibly because both these official runs are poor in the first place. Citri, Berkeley, Xerox and Cornell are independent of most of the others though not necessarily of each other.

Some of these individual results are much in line with the independence observations we would have expected as outlined earlier while some are not. To illustrate this a bit better, in Figure 3 we summarise these observed improvements in retrieval effectiveness from a pairwise fusion by superimposing in red, such improvements on the graph of the perceived conceptual independence of retrieval strategies as observed earlier.



**Fig 3: Conceptual approaches of the pairwise system fusions and observed improvements in retrieval effectiveness from the set of all pairwise fusions**

Figure 3 shows that there is no real consistency in the improvements in effectiveness observed or not when correlated with our *a priori* perceptions of the conceptual independence of retrieval strategies but some improvements are rationalisable. For the reasons due to query expansion outlined above, NMSU and INQ should probably not have been grouped while the Citri/Crnl/Berkeley grouping could have been sub-divided. Thus the fact that 3 of the 10 observed pairwise improvements come from **within** strategy independence groupings, may not be as bad as it initially appears.

From the improvements in pairwise fusions observed, the tabulation of these results suggests that we fuse together triples of inputs in the following combinations where the fusing of the pairwise combinations all improve performance, i.e. cases where fusing A and B yields improvement as does fusing B and C and also fusing C and A.

$$\{GMU, Xrx, Brkly\} \quad \{GMU, Brkly, Citri\} \quad \{GMU, Citri, Crnl\}$$

The fusion of these triples was done and results are presented below

	<b>GMU-Xrx-Brkly</b>	<b>GMU-Brkly-Citri</b>	<b>GMU-Citri-Crnl</b>
<b>P @5</b>	0.536	0.472	0.408
<b>P @10</b>	0.496	0.492	0.404
<b>P @20</b>	0.450	0.412	0.376

**Table 3: Performance Figures for Fusion of selected triples**

The results in Table 3 show that the three-way fusions we tried do yield results which were better than any of the individual and comparable to the performance of the pair-wise constituents, yet these best results were still far less than the performance of the pair-wise fusion of INQ and NMSU.

The final set of experiments we carried out that we report here is to vary the number of documents in the rankings input into the fusion process. Earlier we pointed out that we would fuse document rankings based on the top 100 ranked documents per retrieval strategy. To see if this was enough we fused together all 9 individual rankings into one single document ranking, based on fusing the top 100 per strategy and also based on fusing the top 1000. The results, shown below in Table 4, show no difference for the evaluation measures we use and for that matter for our use of data fusion (precision at rank positions 5, 10 and 20) though in other measures of IR effectiveness such as the interpolated recall-precision averages and average precision over all relevant documents, differences would be seen. It is interesting to observe that fusing all 9 strategies together yields performance which is better than any individual strategy except INQ, but still less than the fusion of INQ and NMSU.

	<b>Fuse all - top 100</b>	<b>Fuse all - top 1000</b>
<b>P @5</b>	0.520	0.520
<b>P @10</b>	0.464	0.464
<b>P @20</b>	0.418	0.418

**Table 4: Performance Figures for Fusion of all 9 strategies**

While the empirical search for the best combination of data fusions could continue we believe that our point has now been made. Fusing together independent retrieval strategies can yield improved retrieval effectiveness with pair-wise or triple-wise data fusion, even at high precision levels as has been observed elsewhere. However the results we have observed are not consistent and expected improvements in some fusion pairs did not materialise.

## 5. Conclusions

The paper has confirmed other results that data fusion can be used to obtain a better level of effectiveness at high precision by carefully choosing information retrieval strategies or agents which are truly different and independent of each other. The data fusion experiments reported here show that data fusion among ranked outputs of independent retrieval strategies works to improve retrieval effectiveness for information retrieval on Spanish texts, however for most of the cases of fusion of rankings, the combined results are still less than that obtainable with a good, effective standalone strategy like INQ. In experiments reported here we tried other combinations of data fusion using three and more retrieval strategies but were unable to obtain improvements in retrieval effectiveness because of the overall noise introduced by having too many retrieval rankings involved, many of which had poor retrieval performance in the first case.

Part of the reason for this disappointment with some fusion results might be that each of the rankings we used was not the best achievable using that particular retrieval strategy. Part of the *modus operandi* of TREC tracks in their early days is that groups submit ad hoc retrieval runs and then post-TREC when a sizable set of relevance judgments are available, perform more runs to vary parameter settings and determine the best combinations. In the case of TREC-4 Spanish, there was very little training data available from previous TRECs and little scope for participants to develop the most effective retrieval settings for their runs. For some groups such as UMass, NMSU and Berkeley quite a bit of work had been input into obtaining better performance while for

others (including ourselves) completion of the task by the deadline was the main goal, with refinement to follow afterwards. The reason why this may affect the interpretation of our results is pointed out by Vogt [7] who examined the question of when it made sense to linearly combine the outputs of different retrieval scores. His summary conclusion is that this is best done “*when both have reasonable performance of similar magnitude, but do not rank documents in a similar fashion*”. This may be the case for some, but not for all of the pairs of runs we used in our fusion experiments.

In the work we have reported here, the data fusion technique we employed has been of the most simple kind. The summation of normalised document scores with each strategy or retrieval agent treated equally presupposes that for a given retrieval situation each retrieval agent is of equal importance. This simplification is useful for cases where the retrieval agents whose outputs are being fused, index document collections which are overlapping but not exactly the same. In other work using data fusion we combine the output of several WWW search engines into one unified ranking yielding more effective and exhaustive retrieval and in this work the fusion is based upon normalised rank positions rather than normalised document scores [11].

Moving beyond the simple applications of data fusion, if some kind of feedback about document relevance, from the user to the system can be employed then this can be used within a given retrieval strategy to re-rank unseen documents, a technique known as relevance feedback and known to improve retrieval effectiveness. Feedback may also be used to adjust the contribution of different retrieval agents to yield what is called “adaptive” data fusion where the fusion process, itself an “agent”, can dynamically adjust the importance and weighting of each retrieval agent in response to its performance in the retrieval session so far, as measure by the number and rank positions of the known relevant documents it has found [12].

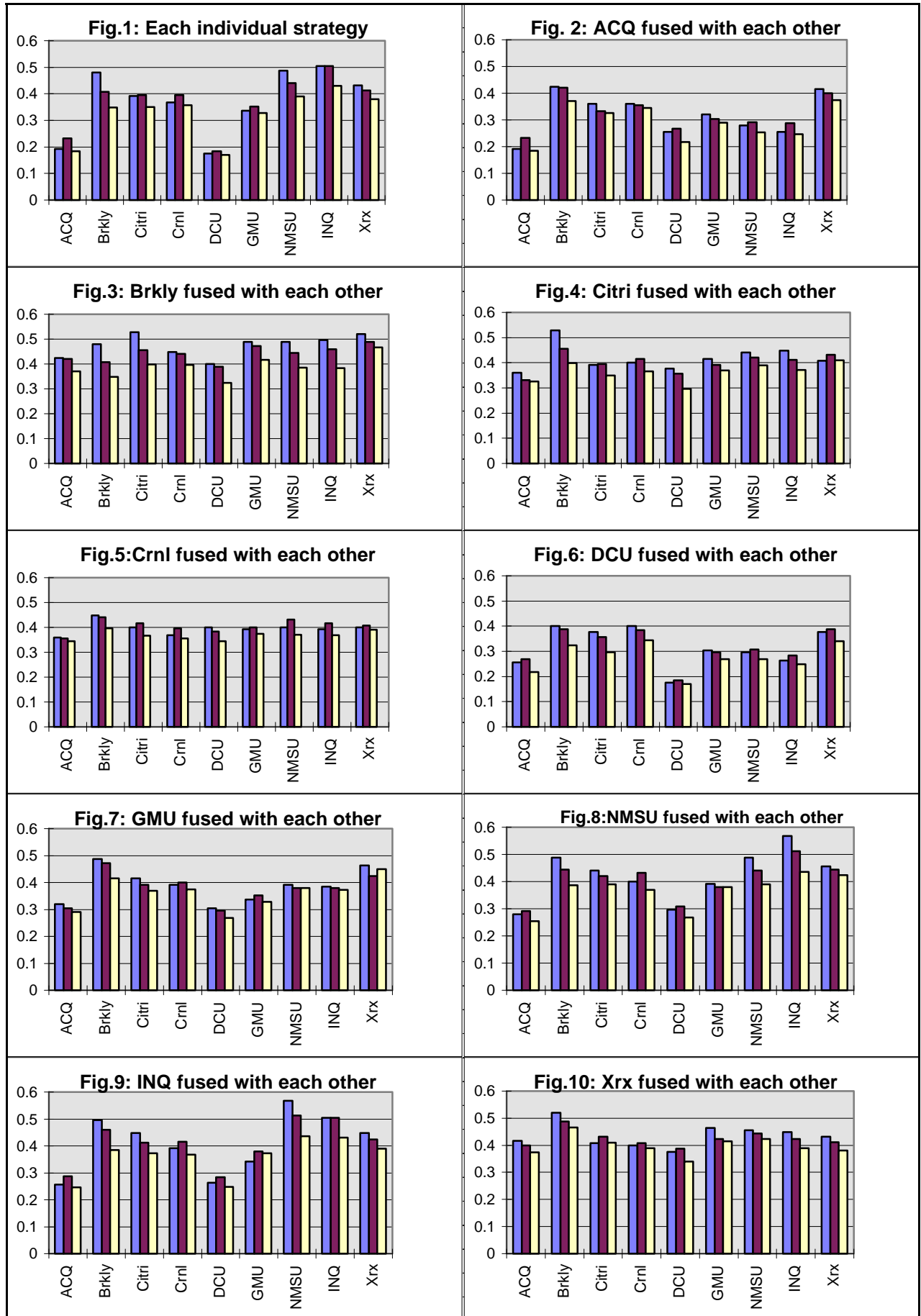
### Acknowledgements:

The author would like to give special thanks to NIST for making the TREC results available and to the two anonymous reviewers whose comments improved this paper.

### References:

1. Brooks RR and Iyengar SS. Robust Distributed Computing and Sensing Algorithm. IEEE Computer 1996; 29(6):53-60
2. Shaw J. and Fox EA. Combination of Multiple Searches. In: Proceedings of TREC-3. Harman D (Ed). National Institute for Standards and Technology Special Publication 500-226, 1995
3. Lee JH. Analysis of Multiple Evidence Combination. In Proceedings of ACM SIGIR Conference, Philadelphia, July 1997, 267-276
4. Belkin NJ, Kantor P, Fox EA and Shaw JA. Combining Evidence of Multiple Query Representation for Information Retrieval. Information Processing & Management, 31(3):431-448, 1995.
5. Croft WB Knowledge Based and Statistical Approaches to Text Retrieval. IEEE Expert 1993; 8(2): 8-11.
6. Vogt CC., Cottrell G., Belew R. and Bartell B. Using Relevance to train a Linear Mixture of Experts. In Proceedings of the Fifth Text Retrieval Conference (TREC-5), National Institute for Standards and Technology Special Publication 500-238, 1997.
7. Vogt CC. When Does it Make Sense to Linearly Combine Relevance Scores ? Poster presented at SIGIR97, Philadelphia 1997. [http://www-cse.ucsd.edu/users/vogt/papers/make\\_sense/make\\_sense.html](http://www-cse.ucsd.edu/users/vogt/papers/make_sense/make_sense.html).
8. Smeaton AF and Harman D. The TREC Experiments And Their Impact On Europe. Journal of Information Science 1997; 23:169-174.
9. Harman D (Ed). Proceedings of TREC-3. National Institute for Standards and Technology Special Publication 500-226, 1995.

10. Harman D (Ed). Proceedings of TREC-4. National Institute for Standards and Technology Special Publication 500-236, 1996.
11. Smeaton AF and Crimmins F. Relevance Feedback and Query Expansion for Searching the Web: A Model for a Digital Library. In: Research and Advanced Technology for Digital Libraries, First European Conference ECDL'97, C Peters and C Thanos (Eds.) Springer LNCS #1324:99-112, 1997.
12. Simonnot B. Active Evaluation and Its Influence on an Adaptive Information Retrieval Process. Presented at the MIRO Workshop, Glasgow, May 1996.



P @ rank

	<b>ACQ</b>	Brkly	Citri	Crnl	DCU	GMU	NMSU	INQ	Xrx
5	0.192	0.424	0.36	0.36	0.256	0.32	0.28	0.256	0.416
10	0.232	0.42	0.332	0.356	0.268	0.304	0.292	0.288	0.4
20	0.184	0.37	0.326	0.344	0.218	0.29	0.254	0.246	0.3733

	ACQ	<b>Brkly</b>	Citri	Crnl	DCU	GMU	NMSU	INQ	Xrx
5	0.424	0.48	0.528	0.448	0.4	0.488	0.488	0.496	0.52
10	0.42	0.408	0.456	0.44	0.388	0.472	0.444	0.46	0.488
20	0.37	0.348	0.398	0.396	0.324	0.416	0.386	0.384	0.466

	ACQ	Brkly	<b>Citri</b>	Crnl	DCU	GMU	NMSU	INQ	Xrx
5	0.36	0.528	0.392	0.4	0.376	0.416	0.44	0.448	0.408
10	0.332	0.456	0.396	0.416	0.356	0.392	0.42	0.412	0.432
20	0.326	0.398	0.35	0.366	0.296	0.37	0.39	0.372	0.41

	ACQ	Brkly	Citri	<b>Crnl</b>	DCU	GMU	NMSU	INQ	Xrx
5	0.36	0.448	0.4	0.368	0.4	0.392	0.4	0.392	0.4
10	0.356	0.44	0.416	0.396	0.384	0.4	0.432	0.416	0.408
20	0.344	0.396	0.366	0.356	0.344	0.374	0.37	0.368	0.39

	ACQ	Brkly	Citri	Crnl	<b>DCU</b>	GMU	NMSU	INQ	Xrx
5	0.256	0.4	0.376	0.4	0.176	0.304	0.296	0.264	0.376
10	0.268	0.388	0.356	0.384	0.184	0.296	0.308	0.284	0.388
20	0.218	0.324	0.296	0.344	0.17	0.268	0.268	0.248	0.34

	ACQ	Brkly	Citri	Crnl	DCU	<b>GMU</b>	NMSU	INQ	Xrx
5	0.32	0.488	0.416	0.392	0.304	0.336	0.392	0.384	0.464
10	0.304	0.472	0.392	0.4	0.296	0.352	0.38	0.38	0.424
20	0.29	0.416	0.37	0.374	0.268	0.328	0.38	0.372	0.45

	ACQ	Brkly	Citri	Crnl	DCU	GMU	<b>NMSU</b>	INQ	Xrx
5	0.28	0.488	0.44	0.4	0.296	0.392	0.488	0.568	0.456
10	0.292	0.444	0.42	0.432	0.308	0.38	0.44	0.512	0.444
20	0.254	0.386	0.39	0.37	0.268	0.38	0.39	0.436	0.424

	ACQ	Brkly	Citri	Crnl	DCU	GMU	NMSU	<b>INQ</b>	Xrx
5	0.256	0.496	0.448	0.392	0.264	0.342	0.568	0.504	0.448
10	0.288	0.46	0.412	0.416	0.284	0.38	0.512	0.504	0.424
20	0.246	0.384	0.372	0.368	0.248	0.372	0.436	0.43	0.39

	ACQ	Brkly	Citri	Crnl	DCU	GMU	NMSU	INQ	<b>Xrx</b>
5	0.416	0.52	0.408	0.4	0.376	0.464	0.456	0.448	0.432
10	0.4	0.488	0.432	0.408	0.388	0.424	0.444	0.424	0.412
20	0.3733	0.466	0.41	0.39	0.34	0.414	0.424	0.39	0.38