

AUGMENTING RELEVANCY SIGNATURES WITH SLOT FILLER DATA

Ellen Riloff and Wendy Lehnert

Department of Computer Science
University of Massachusetts
Amherst, MA 01003

INTRODUCTION

Human readers can reliably identify many relevant texts merely by skimming the texts for domain-specific cues. These quick relevancy judgements require two steps: (1) recognizing an expression that is highly relevant to the given domain, e.g. “were killed” in the domain of terrorism, and (2) verifying that the context surrounding the expression is consistent with the relevancy guidelines for the domain, e.g. “5 soldiers were killed by guerrillas” is not consistent with the terrorism domain since victims of terrorist acts must be civilians¹. The Relevancy Signatures Algorithm attempts to simulate the first step in this process by deriving reliable relevancy cues from a corpus of training texts and using these cues to quickly identify new texts that are highly likely to be relevant. But since this algorithm makes no attempt to look beyond the relevancy cues, it will occasionally misclassify texts when the surrounding context contains additional information that makes the text irrelevant.

As a first attempt to address this problem, we developed a variation of the Relevancy Signatures Algorithm that augments the relevancy signatures with slot filler information. While relevancy signatures classify texts based upon the presence of case frames, augmented relevancy signatures classify texts on the basis of case frame instantiations. Experimental results show that the augmented relevancy signatures can achieve higher precision than relevancy signatures alone while still maintaining significant levels of recall.

AUGMENTED RELEVANCY SIGNATURES

One shortcoming of relevancy signatures is that they do not take advantage of the slot fillers in the concept nodes. For example, consider two similar sentences: (a) “a civilian was killed by guerrillas” and (b) “a soldier was killed by guerrillas”. Both sentences are represented by the same relevancy signature: (killed, \$murder-pass-1) even though sentence (a) describes a terrorist event and sentence (b) does

¹According to the MUC-3 domain guidelines, events that targetted military personnel or installations were not considered to be terrorist in nature.

not. To address this problem, we experimented with augmented relevancy signatures that combine the original relevancy signatures with slot filler information.

Given a set of training texts, we parse each text and save the concept nodes that are generated. For each slot in each concept node², we collect reliability statistics for triples consisting of the concept node type, the slot name, and the semantic feature of the filler.³ For example, consider the sentence: “The mayor was murdered.” The word “murdered” triggers a murder concept node that contains “the mayor” in its *victim* slot. This concept node instantiation yields the slot triple: (murder, victim, ws-government-official). For each slot triple, we then update two statistics: [1] the number of times that it occurred in the training set (N), and [2] the number of times that it occurred in a relevant text (N_R). The ratio of N_R over N gives us a “reliability” measure. For example, .75 means that 75% of the instances of the triple appeared in relevant texts.

Using these statistics, we then extract a set of “reliable” slot triples by choosing two values: a reliability threshold R_{slot} and a minimum number of occurrences threshold M_{slot}. These parameters are analogous to the relevancy signature thresholds. The triples that satisfy the reliability criteria become our set of “reliable” slot filler triples.

The algorithm for classifying texts is fairly simple. Given a new text, we parse the text and save the concept nodes that are produced during the parse, along with the words that triggered them. For each concept node, we generate a (triggering word, concept node) pair and a set of slot triples. If the (triggering word, concept node) pair is in our list of relevancy signatures, and the concept node contains a reliable slot triple then we classify the text as relevant. If not, then the text is deemed irrelevant. Intuitively, a text is classified as relevant only if it contains a strong relevancy cue and the concept node enabled by this cue contains at

²We only collect statistics for top-down slots, i.e. slots that were predicted by the concept node.

³Since slot fillers can have multiple semantic features, we create one triple for each feature. For example, if a murder concept node contains a victim with semantic features *ws-human & ws-military* then we create two triples: (murder, victim, *ws-human*) and (murder, victim, *ws-military*).

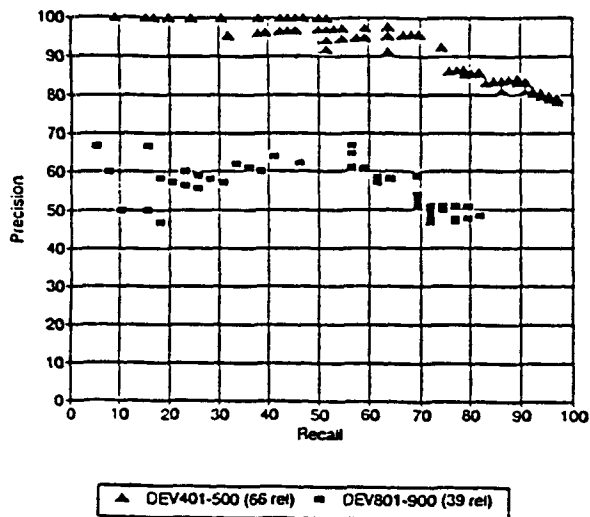


Figure 1: Relevancy Discriminations on Two Separate Test Sets Using Relevancy Signatures

least one slot filler that is also highly correlated with relevance.

COMPARATIVE EXPERIMENTS

We compared the performance of the augmented relevancy signatures with the original Relevancy Signatures Algorithm in order to measure the impact of the slot filler data. We tested the augmented relevancy signatures on the same two test sets that we had isolated for our original experiments, after training on the remaining 1300 texts. Figure 1 shows the original results produced by the Relevancy Signatures Algorithm and Figure 2 shows the results produced by the augmented relevancy signatures. Each data point represents a different combination of parameter values.

These graphs clearly show that the augmented relevancy signatures perform at least as well as the original relevancy signatures on these two test sets. The most striking difference is the improved precision obtained for DEV 801-900. There are two important things to notice about Figure 2. First, we are able to obtain extremely high precision at low recall values, e.g., 8% recall with 100% precision and 23% recall with 90% precision. Relevancy signatures alone do not achieve precision greater than 67% for this test set at any recall level. Second, although there is a very scattered distribution of data points at the lower recall end, we see consistently better precision coupled with the higher recall values. This trend suggests that the augmented relevancy signatures perform at least as well as the original relevancy signatures when they are working with statistically significant numbers of texts.

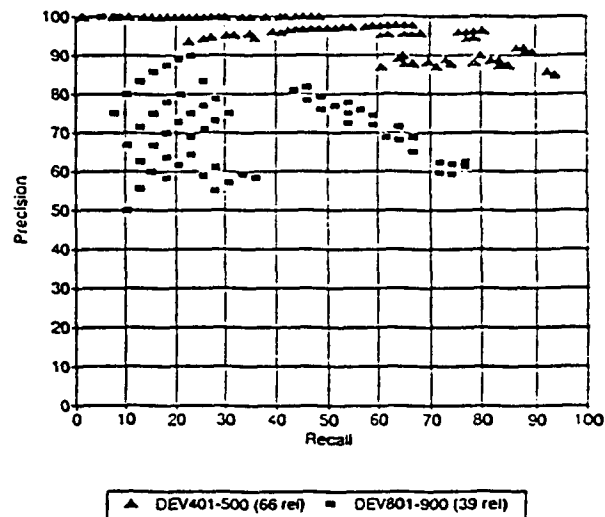


Figure 2: Relevancy Discriminations on Two Separate Test Sets Using Augmented Relevancy Signatures

Furthermore, the Relevancy Signatures Algorithm demonstrated extremely strong performance on DEV 401-500 and it is reassuring to see that the augmented relevancy signatures achieve similar results, perhaps even showing a slight improvement at the higher recall values. The highest recall level obtained with extremely high precision by the original relevancy signatures was 67% with 98% precision. The augmented relevancy signatures achieved significantly higher recall with nearly the same precision, 77% recall with 96% precision.

CONCLUSIONS

We have demonstrated that augmented relevancy signatures can achieve higher levels of precision than relevancy signatures alone while maintaining significant levels of recall. Augmenting relevancy signatures with slot filler information allows us to make more fine-grained domain relevancy classifications. Furthermore, the additional slot filler data can be acquired automatically from a training corpus using the same selective concept extraction techniques needed to collect the relevancy signatures. Combining slot filler information with relevancy signatures is a promising approach for improving precision without sacrificing significant recall in text classification tasks.