14th International Conference on Information Fusion
Chicago, Illinois, USA, July 5-8, 2011

Evaluating Spreading Activation for Soft Information Fusion

Michael Kandefer and Stuart C. Shapiro
Department of Computer Science and Engineering
and National Center for Multisource Information Fusion
State University of New York at Buffalo
201 Bell Hall, Buffalo, New York 14260
{mwk3,shapiro}@buffalo.edu

***Abstract*—A soft-information fusion process produces refined estimates of soft-information, such as natural language messages. Information resulting from a soft-information process can be used to retrieve related, relevant information from background (a-priori) knowledge sources using contextual "cues" contained in those messages, a process we call "Context-Based Information Retrieval (CBIR)". These retrieval results can be used to aid further understanding, and other fusion operations (e.g., data association). CBIR process performance is dependent on the choice of algorithms and parameters for those algorithms, and it is crucial that these are chosen appropriately for the problem domain the CBIR algorithm is used to aid. In this paper an *f-measure* evaluation of two spreading activation algorithms and their parameters is given using a soft information fusion process in a counterinsurgency domain. This evaluation takes place in two phases. The first phase executes the algorithms over a range of values in order to determine how those parameters affect the performance of the algorithms, and to set these parameters for future use. The second phase compares the results of these algorithms using the parameter settings learned from the first phase.**

Keywords: human-based sensing, soft information fusion, context exploitation

Introduction

Information fusion is defined as the process of associating, correlating, and combining data and information from single and multiple sources to achieve refined estimates of characteristics, events, and behaviours for observed entities in an observed field of view [1]. Recently there has been an increased interest in applying information fusion to soft information sources such as natural language ([2], [3], [4], [5]). The goal of such research is to fuse information contained within natural language data with more refined data in order to establish a better understanding of the domain in question. A system that is able to properly accept and use natural language information should be capable of evaluating the message in the context of its background knowledge, state estimate, and other—possibly conflicting—messages [4].

One solution suggested to provide this capability for similar problems [6] and for soft-information fusion [4] is the use of Context-Based Information Retrieval (CBIR), which is a process used to retrieve relevant information from background knowledge sources using contextual "cues" contained in some input parameter. For a soft-information fusion process these "cues" would be found in the natural language messages. By using the CBIR process other natural language and fusion processes can exploit the relevant information retrieved to perform:

- Co-reference resolution by using the retrieved information to deduce the likely reference for the natural language term that has multiple references. For example, in the domain the word "defense" can refer to several possible ontological entities. Contextual information can be useful in deducing the referent.
- Data association by using the retrieved information to help deduce two message entities are the same, and thus, require merging.

CBIR process performance is dependent on the choice of algorithms and parameters for those algorithms, and it is crucial that these are chosen appropriately for the problem domain the CBIR algorithm is used to aid. Previous work on the subject has introduced a method for evaluating CBIR algorithms [6]. This paper continues this work by using that methodology to compare two spreading activation algorithms in a soft-information fusion domain. The algorithms are:

- Texai Algorithm - A standard spreading activation algorithm [7], and
- ACT-R Declarative Memory Activation Algorithm - A spreading activation algorithm developed for the ACT-R cognitive model [8].

The Texai spreading activation technique was chosen for the evaluation because it was designed as an information retrieval technique for propositional graphs. Since we are using propositional graphs for the representation of information in this problem space and require an information retrieval technique the algorithm was chosen. The ACT-R activation algorithm was chosen because it is frequently used as an example of spreading activation techniques in cognitive psychology textbooks ([9], [10]). ACT-R is also a well known cognitive architecture. However, ACT-R's spreading activation technique was not designed for information retrieval, but for ranking information in a general search. This paper will discuss why this will not hamper the technique as an information retrieval tool, but because it was not designed for the purpose of information retrieval we expected Texai to be a better technique for our purposes.

This evaluation takes place in two phases. The first phase is a type of learning phase that establishes useful parameters for each algorithm for retrieving information in the domain. This phase shows how one can set parameters in spreading activation algorithms using the *distance from the optimal* measure. The second phase compares the results of these algorithms using these learned parameters against the message set used to learn them.

Context-Based Information Retrieval for Soft-Information Fusion

One problem of interest for soft information fusion is exploiting natural language understanding in counterinsurgency (COIN) ([4], [11]). In the COIN domain the goal is to convert natural language messages into propositional graph representations for future processing and reasoning. Messages are small English paragraphs containing information about activity that might be of interest to a COIN operative. An example message from the Soft Target Exploitation and Fusion (STEF) research project data set [12] is message 2, which says:

> Source said a Sunni youth he knows to be about 20 years old, Khalid Sattar, has become increasingly vocal in denouncing the U.S. at several mosques in Adhamiya.

In these domains it is helpful to place a message in context by retrieving additional information about the message entities, such as the geographical entities and people, from background knowledge sources, and then reasoning about the message and the retrieved information to produce a "contextually enhanced message". The contextual enhancement process relies on two components, a Context-Based Information Retrieval (CBIR) process, and forward-inference reasoning process. These processes are depicted in Fig. 1.

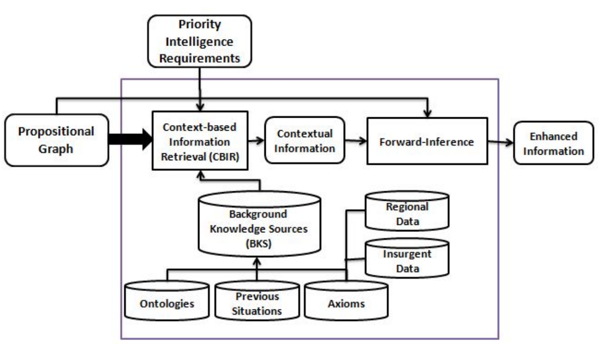

Figure 1: Context-Based Information Retrieval Process

A general CBIR process receives input (*I*), which contains the *contextual constraints* and other information about the situation; and the background knowledge (*BKS*) containing any knowledge that will be evaluated by the CBIR procedure. With this the CBIR procedure produces a subset of the background knowledge, called the *retrieved propositions*, which can then be combined with *I*, and given to a reasoning engine to perform forward inference on.

In the COIN domain the *I* is a propositional graph representing the message and a set of *Prioritized Intelligence Requirements* (PIR), which are a means of providing a focus for the counterinsurgency operation (e.g., what is the goal of the operation). The *BKS* include:

- **Previous Messages**: Information stored from previously processed messages, including the enhanced propositional graph representations of those messages, the authors, and timestamps.
- **Ontologies**: Categorical relationships among entities in the domain, as well as relationships that can hold between them.
- **Axioms**: Additional axioms for reasoning about the domain (e.g., threat assessment).
- **Insurgent Data**: Information about known insurgents and activity (e.g., affiliations, contacts, pseudonyms), and rules for reasoning about counterinsurgency.
- **Regional Data**: Information about the regions of the world (e.g., geospatial location, population, alternative names) and how to reason about regions (e.g., connected regions, proximity).

The forward-inference process performs forward-inference with *I* and the *retrieved propositions*, which include rules for reasoning, to generate conclusions from the combined information. After the inferences are finished the results are merged with the *propositional graph*. This combined propositional graph is called the *contextually enhanced graph*.

Crucial to producing the *contextually enhanced graph* output is the success of the CBIR process. If it does not retrieve enough information then the reasoner will not conclude everything it can about the entities in the message. Such information could be crucial to a COIN operative. If the CBIR procedure brings in too much information the reasoner could spend too much time performing operations that don't end up producing a conclusion, costing computational resources.

Spreading Activation Algorithms

General Spreading Activation and Propositional Graphs

Spreading activation is an information retrieval procedure that was developed for propositional graphs, and based on models of cognition ([8], [13], [14]). Propositional graphs have several properties that are useful in the domain and information fusion in general. Since propositional graphs are types of semantic networks [15], [16], [17], they have all of the advantages of semantic networks, like RDF, for soft information fusion [18], [4]. In addition to these advantages, propositional graphs have great expressivity in their ability to represent complex assertions. Propositional graphs have the following capabilities:

- Propositions are represented as terms in the representational language, thus propositions can be the arguments of other propositions (e.g., relationships).
- Relationships are *n-ary* (i.e., they can have any number of arguments). In semantic networks relationships are binary, but propositional networks lack this restriction [16]. Other uses of n-ary relations in graphs for fusion applications are discussed in [19].

As a result of these advantages we have used propositional graphs for representing the messages and background knowledge sources, and use the SNePS 3 knowledge representation system [20] for the implementation of the representation. An example of a SNePS 3 propositional graph representing message 2 is shown in Figure 2.

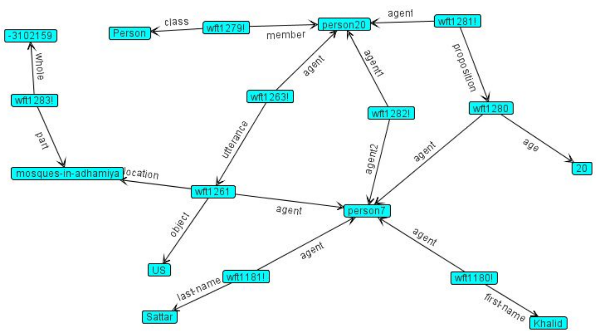

Figure 2: A SNePS 3 propositional graph of STEF message 2

SNePS 3 represents the propositional message content as uniquely identified *well-formed-term (wft)* nodes. Each node is given its meaning by the arcs emanating from it, the nodes those arcs point to, and other structures of the SNePS 3 knowledge representation system not shown in the graph (e.g., case-frames). The arcs represent the arguments of a logical predicate function,[1] and the nodes pointed to are the fillers. For example, *wft1282* represents the proposition that: "person20 (the source of the message) knows person7 (Khalid Sattar)". It gets this meaning from the "knows" case-frame which indicates that *wft* nodes that have only an *agent1* and *agent2* arc emanating from them represent that the node pointed to by *agent1* knows the node pointed to by *agent2*. The *person7* and *person20* nodes are called "base nodes". Base nodes don't have any meaning in and of themselves, but get their meaning from their relationships with other nodes. The knowledge representation system knows that *person7* is named "Khalid Sattar" only because it has a relationship between *person7* and the first name "Khalid" and last name "Sattar" (i.e., *wft1180* and *wft1181*). A *wft* with an exclamation point after it indicates that it is asserted in the knowledge-base. Those *wfts* without exclamations are unasserted, but can still be utilized as arguments to other *wfts*. For example, *wft1263* is asserted and has as an argument the unasserted *wft1261*. The former represents the act of the saying something ("The source said..."). The latter represents what was said ("Khalid Sattar was denouncing the US outside mosques in Adhamiya".)

Since we are using propositional graphs for the representation of information in the domain it was practical to explore retrieval operations that were designed for propositional graphs, like spreading activation. Spreading activation works by retrieving those nodes in a graph that achieve a specified activation value using this general algorithm:

[1]Predicate names are not always represented in the graphs, but are always stored as case-frame data structures in SNePS 3.

1) Initiate a pulse into the propositional graph at specified nodes. This is called a "cue" and sets these nodes to have an activation value higher than the threshold.[2]
2) Spread this pulse throughout the graph while it has energy to activate nodes.
3) The pulse dies when the activation energy runs out, and the algorithm terminates.
4) Activated proposition nodes are considered retrieved for use in future processing, like reasoning.

The key step in this process is determining whether a node is active or not (Step 2), which allows the algorithm to spread the activation throughout the graph. This is accomplished through an *activation function*, which calculates the *activation level* of a node and compares this value to an *activation threshold* parameter. If the value is above the threshold the node is considered active. If no more nodes can be activated, the algorithm terminates, this is called the *terminating condition*. *Activation functions* and *terminating conditions* vary between algorithms, and all spreading activation algorithms permit setting parameters, like the *activation threshold*. In the algorithms covered in this paper *activation levels* are real values in the closed interval between 0.0 and 1.0. The *activation threshold* parameter is crucial to the performance of the algorithms, and determines how much information is retrieved, as such it is a crucial variable to the evaluation. Other crucial variables (and those less so), will be discussed as encountered.

Fine tuning the algorithms to retrieve enough information to be useful without retrieving too much is crucial to successful use of the process. Figure 3 shows an example of spreading activation results. The figure shows the propositional graph with retrieved information added after using the Texai algorithm (discussed below) and message 2 as a "cue". The algorithm spread outward conservatively with the settings given[3] and retrieved from the background knowledge sources three propositions:

- *wft1179*: "Khalid Sattar is a Person."
- *wft1262*: "Denouncing the U.S. is an indicator of terrorist activity."
- *wft1266*: "If someone is engaged in a type of terrorist activity, they may be an insurgent."[4]

Texai Algorithm

The Texai knowledge representation and natural language processing system [7] uses a spreading activation algorithm on a propositional graph. The Texai system uses a simple activation function to calculate the activation of a node (A'_j) given the set of adjacent nodes (N):

$$A'_j = A_j + \sum_{i \in N} A_i * W_{ij} * D$$

[2]A value of 1.0 is used for the initial activation value since it is recommended in the Texai approach, and will always be greater than or equal to the *activation threshold*.

[3]The *activation threshold* was 0.5 and *decay* was 0.9 (c.f., *Texai Algorithm*).

[4]The actual representation of this reasoning rule is much more specific. As of publication SNePS 3 is currently in development and does not support yet the more robust reasoning capabilities of its predecessor.

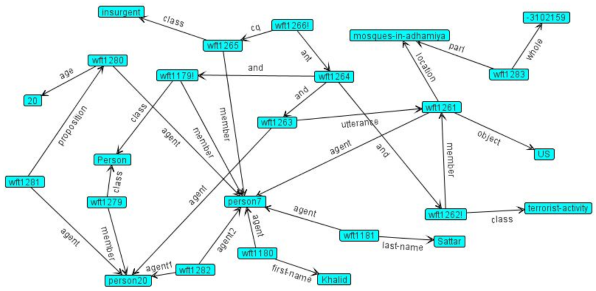

Figure 3: A SNePS 3 propositional graph after Texai spreading activation

The *activation function* takes into account any previous activation the node may have received (A_j), the activation of adjacent nodes (A_i), the weights of the edges between nodes (W_{ij}), and a decay factor (D). The Texai spreading activation algorithm does not specify a means of calculating weights, but does specify that they are values in the closed interval between 0.0 and 1.0. An approach similar to ACT-R will be used. The technique uses the degree of the node to calculate weighting between nodes. Each arc connected to the node *i* is given the value of 1.0 divided by the number of arcs pointing to the node. In Texai, the *decay factor*, a value also in the closed interval between 0.0 and 1.0, is used to impact how quickly the pulse decays as it spreads through the graph, it is not to be confused with the concept of decay as a constraint [14] in constrained spreading activation algorithms. The Texai *decay factor* is a variable of interest for evaluation. Texai uses the same *terminating condition* as the general algorithm given previously (i.e., the nodes recently spread into don't exceed the *activation threshold*). With these parameters and ranges explained the Texai *activation function* evaluated is:

$$A'_j = A_j + \sum_{i \in N} A_i * (1/|N|)) * D$$

ACT-R Declarative Memory Activation Algorithm

ACT-R is a modular cognitive architecture that models cognitive process of humans [21] and its *declarative memory module* is a frequent example of spreading activation in cognitive psychology textbooks [9]. ACT-R's modules communicate information to a central production system, that in turn places information in their buffers to determine how they operate. When the production system places information into the declarative buffer the declarative memory module uses the information as a "cue" into a spreading activation algorithm on a propositional graph to retrieve one memory "chunk" (i.e., proposition) that best matches the "cue" after ranking all of the "chunks" with their *activation value*. Despite only using it to retrieve one chunk, the activation calculation can still be used as a means of retrieving multiple chunks by using it as the *activation function* in the general spreading activation algorithm.

The spreading activation algorithm for ACT-R is specified in ([8], [22]). The specification does not provide a *terminating condition* for using the algorithm as a means of information retrieval in large-scale knowledge bases.[5] To alleviate this, we used the general spreading activation *terminating condition*.

Activation Equation: The ACT-R spreading activation algorithm requires a set of nodes that will contribute to the activation of other nodes, called the "context" (C). Initially these will be the nodes contained in the pulse. If more nodes fire as a result of the activation calculation, they will become to new "context". This process repeats until termination. ACT-R uses the following *activation function* for calculating the *activation level* of a node (A_i) given initial "context" nodes *C*:

$$A_i = B_i + \sum_{j \in C} W_j * S_j$$

To calculate the activation of the node (i) the *activation function* takes into account the *base-level activation* (B_i), the *attentional weighting* of "context" nodes *C*, and the *associative strength* (S_j) of the connection between nodes in the "context" *C* and i. B_i, W_j, and S_j are explained more fully below.

Base-level Activation Equation: The *base-level activation* models the effect of past usage of a memory "chunk" and how those uses influence the retrieval of that "chunk" later. The following is used to calculate *base-level activation*:

$$B_i = ln(\sum_{k=1}^{n} t_k^{-d})$$

Here the t_k indicates the time since the kth usage of the term (out of n usages). Neither the ACT-R specification nor the manual describe how times are stored. In the implementation used for evaluation we assume a new cycle (i.e., every time a pulse is encountered) indicates the passage of one time increment and all prior knowledge in the background knowledge sources occurred at the first time period. A pulse is used to add new times encountered to those "chunks" contained in the pulse. These assumptions are listed for completeness and replicablity, but do not offer a variable of interest for the evaluation since the evaluation evaluates pulses as retrieval queries from system start up and no new timestamps will be provided for the "chunks" as a result. The d indicates a *decay* parameter, a positive real value[6] that determines how much influence time passage has on the base-level activation. The *decay* has negligible influence on the algorithm performance in the evaluation since no new timestamps are created.

Attentional Weighting Equation: The *attentional weighting* models a form of context sensitivity in memory retrieval (i.e., that certain "cues" contribute more to retrieval than others). This is done by associating the "chunks" in the "context"

[5] ACT-R representations typically use smaller knowledge bases than those in large-scale systems, like Cyc [23], that require information retrieval techniques. As such, the activation calculation used in ACT-R gives a ranking to all the information in declarative memory and then selects the best ranked results as a match.

[6] The ACT-R specification [8] recommends a value of 0.5 for d after numerous tests, but this was for the retrieval of one chunk and may be different for using the spreading activation algorithm for information retrieval.

with the individual ACT-R buffers that they originated from and calculating their weight as a percentage of the number of other chunks that originate from that buffer. To calculate the *attentional weighting* (W_j) the following is used:

$$W_j = 1/n$$

Where n is the number of nodes that come from the same buffer as j. However, in the implementation used for our evaluation we do not have buffers like ACT-R, and thus all nodes will be considered as originating from the same buffer. Because of this assumption all nodes in the "context" contribute equally to the *attentional weighting* factor, and n is the *cardinality* of **C**, $|C|$.

Associative Strength Equation: The *associative strength* is used to model how strong the associations are between the "context" nodes and i. To calculate the *associative strength* a heuristic from the ACT-R manual [22] is used instead of the underspecified equation discussed in [8].[7] The equation from [22] is as follows:

$$S_j = S - ln(fan(j))$$

Here S is the ***maximum associative strength*** parameter, a value that can be any real number, and is a variable of interest for evaluation. The $fan(j)$ is the number of arcs in the propositional graph connected to j.

With the parameters and assumptions explained the ACT-R *activation function* evaluated is:

$$A_i = ln(\sum_{k=1}^{n} t_k^{-0.5}) + \sum_{j \in C}(1/|C|) * (S - ln(degree(j)))$$

METHODOLOGY

The evaluation of the spreading activation algorithms in the COIN domain took part in two phases. In **Phase I** we examined how changes in the parameters affect the performance of the algorithms. In **Phase II** we compared the algorithms with their best parameter settings using the information gathered from the first phase. To evaluate the algorithms we use the following:

- A propositional graph knowledge representation and reasoning (KRR) system, SNePS 3 [20],
- The two algorithms implemented to work with the SNePS 3 graph structure,
- Four messages from from the STEF dataset represented in SNePS 3 to be used as "cues" into the spreading activation algorithms,
- A subset of the National Geospatial-Intelligence Agency: GEOnet Names Server (NGA: GNS) [24] and hand crafted background information about people in the domain represented in SNePS 3. These represent the background knowledge sources (BKS) for the domain, and
- Seven hand crafted reasoning rules for reasoning about the COIN domain that rely on information from the messages and the BKS. These are also considered part of the BKS.[8]

[7]The equation provided in [8] calculates the *associative strength* as $S_{ji} \approx ln(prob(i|j)/prob(i))$, but provides no specification for calculating the probabilities in a propositional graph.

Texai Algorithm		
Variable	Range	Increment
Activation Threshold	0.0 - 1.0	0.1
Decay of Pulse (D)	0.0 - 1.0	0.1
ACT-R Declarative Memory Algorithm		
Variable	Range	Increment
Activation Threshold	0.0 - 0.19	0.01
Maximum Associative Strength (S)	0.5 - 5.0	0.5

Table I: Variables and Ranges Evaluated

The four messages were:

- Message 1: Approximately 40 worshippers outside the Imam Shaykh Hadid shrine in Adhamiya expressed increased hostile sentiment against U.S. troops yesterday.
- Message 2: Source said a Sunni youth he knows to be about 20 years old, Khalid Sattar, has become increasingly vocal in denouncing the U.S. at 4 or 5 mosques in Adhamiya
- Message 3: Bookstore owner on Dhubat Street in Adhamiya said about a dozen customers are asking if he has any books, magazines, or other material on al-Qaeda.
- Message 4: Large gathering of 20 to 30 teenagers ages 15-19 chanted anti-U.S. slogans outside the Jami' al Kazimiyah mosque in Adhamiya.

Phase I consisted of sampling the performance of the spreading activation algorithms on these messages over a range of values to approximate the best settings for the various parameters in the COIN domain. The process is as follows: (1) Iterate over the variable settings using the selected range and increments, (2) A single message will be given to the algorithm and the results evaluated, (3) The system will then be reset, and the next message will be evaluated, and (4) Repeat the previous steps for all the parameter settings in the iteration. The variables of interest and ranges used in the evaluation are given in Table 1.

We chose the Texai values and ranges because these are the ranges these values can take on as specified for the *Texai* algorithm [7]. The ACT-R values were chosen after preliminary testing demonstrated that the algorithm did not return any results outside of this range.

To score the variable sampling of the two spreading activation techniques we used the accepted practice for evaluating information retrieval results, the calculation of an *f-measure* [25]. An *f-measure* score is between 0.0 and 1.0, with 0.0 indicating the poorest result and 1.0 a perfect retrieval. The calculation of an *f-measure* requires a set of *retrieved propositions*, which will be the results of the spreading activation algorithms in this evaluation, and a set of *relevant propositions*.

The set of *relevant propositions* represent what the desired results should be. For our evaluation we used a technique

[8]The SNePS 3 KRR system, background knowledge sources and means of loading them into SNePS 3, message representations, and code for evaluating the algorithms is available at http://www.cse.buffalo.edu/~mwk3/Papers/evaluation.html.

Recall (r)	**Precision (p)**	**F-measure (F)**
$r = \frac{\lvert\{relevant\ propositions\} \cap \{retrieved\ propositions\}\rvert}{\lvert\{relevant\ propositions\}\rvert}$	$p = \frac{\lvert\{relevant\ propositions\} \cap \{retrieved\ propositions\}\rvert}{\lvert\{retrieved\ propositions\}\rvert}$	$F(r,p) = \frac{2rp}{r+p}$

Figure 4: Formulas for computing the *f-measure* [25]

based on the *distance from the optimal* [6]. This technique uses the input "cue" (*I*) and the contents of the background knowledge sources (*BKS*) to determine which portion of the *BKS* will be used as part of an inference (forward or backward). Those propositions that contribute to any inferences made, excluding those from *I*, are considered the *relevant propositions*, since they are the only propositions from the *BKS* necessary for drawing the conclusions we want in the domain. This method of establishing relevancy is useful for a domain that uses reasoning, like the COIN domain, since it will necessarily indicate as relevant all the information needed to reason about a given input. However, some domains may consider relevant information to be more than that which is used in the reasoning processes. For example, it is possible that a suspect may be a known insurgent in the background knowledge sources, but if no reasoning rule uses that information the information will not be considered relevant by the distance from the optimal technique. The *distance from the optimal* offers an objective means of establishing what is relevant from the background knowledge sources, but does not take into account the subjective opinions of COIN experts.

The *distance from the optimal* technique described by Kandefer and Shapiro [6] also requires a reasoner capable of maintaining origin sets,[9] and performing reasoning to populate the origin sets of some query proposition, and using that as the *relevant propositions*. However, SNePS 3 currently lacks origin sets and the number of rules were small, so we chose to determine manually what the origin sets are for each "cue" encountered.

With the *relevant propositions* established for each message the *f-measure* can be calculated. An *f-measure* is the harmonic-mean of two other calculations (not shown in the results). The first calculation is the *recall*. The *recall* is also a value between 0.0 and 1.0 and is the fraction of the *relevant propositions* that are included in the *retrieved propositions*. The second calculation is the *precision*. The *precision* is the fraction of the *retrieved propositions* that are *relevant propositions*. These calculations are depicted in Figure 4.

To learn the best settings for the algorithms we calculated the *f-measure* for each algorithm on the four messages using the parameter ranges and increments in Table 1. Since two parameters were evaluated per algorithm, this resulted in four *f-measure* matrices per algorithm. We took the mean of the four matrices and then selected the cell in the resulting matrix with the highest *f-measure*. The parameter settings corresponding to this cell were used as the "best" settings for **Phase II**. For Texai the best parameter settings were an *activation threshold* of 0.5 and decay of 0.9. For ACT-R this was an *activation threshold* of 0.04 and *maximum associative strength* of 2.0.

Phase II uses the results of best settings learned from **Phase I** to compare the two algorithms with these settings. This was done to test how well the algorithms can be trained using an initial sample set, and comparing their performance using the learned settings. Since only four messages were hand crafted, the four messages are used again for the comparison.

Evaluation Results

Phase I - Parameter Learning

Texai Algorithm

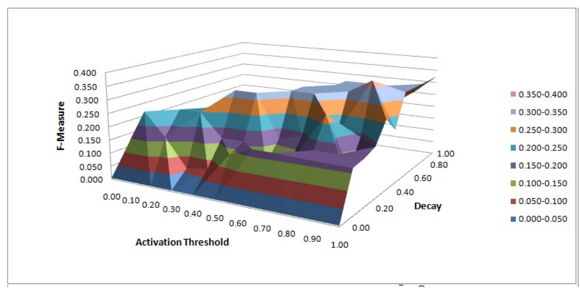

Figure 5: Average F-Measures for Texai

Figure 5 shows the average *f-measure* results of varying the *activation threshold* from 0.0 to 1.0 and the *decay* from 0.0 to 1.0 using the Texai spreading activation algorithm on the four messages. The average maximum *f-measure* of 0.375 occurred when the *activation threshold* was 0.5 and the *decay* was 0.9.

ACT-R Declarative Memory Activation Algorithm

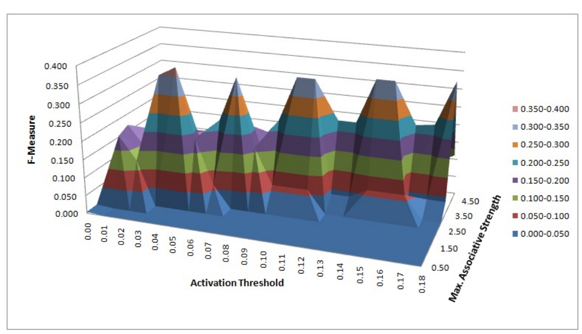

Figure 6: Average F-Measures for ACT-R Declarative Memory

Figure 6 shows the results of varying the *activation threshold* from 0.0 to 0.19 and the *maximum associative strength*

[9] An origin set for a proposition is the set of propositions used in the derivation of that proposition. Origin sets originate from *relevance logic* proof theory [26].

from 0.5 to 5.0 using the ACT-R declarative memory activation function on the four messages. The maximum average *f-measure* of 0.375 occurred when the *activation threshold* was set to 0.04 and the *maximum associative strength* was set to 2.0.

Phase II - Comparison

Message	Texai	ACT-R
Message 1	0.5	0.0
Message 2	1.0	0.75
Message 3	0.0	0.5
Message 4	0.0	0.25
Mean	0.375	0.375
Standard Deviation	0.41	0.28

Table II: F-Measures for the Predicted Parameter Settings

The results of using the best average settings for the two algorithms on the four messages are shown in Table II. For Texai the *activation threshold* was set to 0.5 and the *decay* was set to 0.9. For ACT-R the *activation threshold* was set to 0.04 and *maximum associative strength* to 2.0. These were chosen since they were where the maximum average *f-measure* occurred as depicted in Fig. 2–3, and are the best settings using the method discussed in the **Methodology** section.

DISCUSSION

Phase I was successful in establishing an understanding of how the variables of interest influenced the spreading activation algorithms in question. In the Texai spreading activation algorithm the *activation threshold* had it's best performance when set to 0.5 (Fig. 5). It was at this point that there was a balance between the *precision* and *recall*. As the threshold increased the *recall* was worse; when it was decreased *precision* was worse. The *decay* tended to cause poorer performance as its value decreased (meaning it had greater impact on the spread of the pulse). As such, higher values tended to result in better *recall* without influencing *precision* greatly, which results in better *f-measures*.

In the ACT-R declarative memory spreading activation algorithm the *activation threshold* functioned similarly to Texai. The algorithm tended to generate it's best results at an average *activation threshold* of 0.04 and a *maximum associative strength* of 2.0 (Fig. 6). Like with Texai these values offered the best balance of *precision* and *recall*, though *precision* was frequently poor. Unlike Texai's *decay*, deviations from 2.0 would cause sharp deceases in the *f-measure* unless the threshold was also adjusted. Though a maximum f-measure average was found among the four messages, ACT-R's declarative memory module can perform successfully under multiple *maximum associative strength* settings (as seen by the numerous peaks), so long as the spreading activation threshold is adjusted as well.

The results of **Phase II** show that when using the best average parameter values learned from **Phase I** that the ACT-R declarative memory module and Texai's spreading activation algorithm have the same average performances. However, ACT-R's declarative memory algorithm has a lower standard deviation, suggesting that it's settings would have performance more closer to the shared average performance with other inputs. This difference was mostly due to poor performance with messages 3 and 4 by Texai. Messages 3 and 4 did not differ considerably from the other messages representationally and all messages required different reasoning rules, but approximately the same amount. These results also show that even with a few messages a learning phase can be used to generate parameters that result in good performance for the algorithms (an *f-measure* as high as 1.0) in this domain. If more messages were used using a similar propositional network representation it is not expected that they would change the results of this evaluation (particularly for ACT-R), so these four messages serve as a small, but useful predictive set for the domain.

CONCLUSIONS AND FUTURE WORK

Several evaluations of using spreading activation techniques for context-based information retrieval in a counter insurgency (COIN) domain were given. The two methods were the Texai spreading activation function and ACT-R declarative memory activation function. These evaluations showed:

- Both ACT-R and Texai offer similar average performance in this domain given the data we have available to us,
- The ACT-R algorithm offers the best predictive performance in this domain given the data we have available to us, and
- That the ACT-R algorithm is suitable for information retrieval, and not just result ranking.

Though much was learned about the spreading activation parameters from the evaluation in **Phase I**, and **Phase II** showed how parameter settings learned can be used to set the parameters for future use of the spreading activation algorithms in the COIN domain, our future work on using CBIR for soft information processing will focus on:

- Creating uniformity between the message representations, which are currently hand-crafted, by implementing an automated propositionalizer that takes in natural language messages and converts them into a propositional graph, as specified in [18], [4].
- Expanding the number of messages evaluated by examining and encoding into SNePS 3 networks messages from a dataset created internally. This dataset will better reflect a COIN operation, and handle limitations encountered in the STEF dataset.
- Development of more reasoning rules since they are the crux of determining the *relevant propositions* using the *distance from the optimal* approach.
- Evaluation of ACT-R's base-level-learning. To examine this aspect of the ACT-R algorithm the domain will be examined as a running process, where previous messages encountered impact the base-level activation of nodes.

ACKNOWLEDGEMENTS

This research activity has been supported in part by a Multidisciplinary University Research Initiative (MURI) grant (Number W911NF-09-1-0392) for Unified Research on Network-based Hard/Soft Information Fusion, issued by the US Army Research Office (ARO) under the program management of Dr. John Lavery

References

[1] D. L. Hall and J. Llinas, "An introduction to multisensor data fusion," in *Proceedings of the IEEE*, vol. 85(1), pp. 6–23, January 1997.

[2] D. L. Hall, J. Llinas, M. McNeese, and T. Mullen, "A framework for dynamic hard/soft fusion," in *Proc. 11th Int. Conf. on Information Fusion, Cologne, Germany*, 2008.

[3] M. A. Pravia, O. Babko-Malaya, M. K. Schneider, J. V. White, C.-Y. Chong, and A. S. Willsky, "Lessons learned in the creation of a data set for hard/soft information fusion," in *12th International Conference on Information Fusion*, 2009.

[4] M. Kandefer, M. Prentice, and S. C. Shapiro, "Tractor: A framework for soft information fusion,," in *Proceedings of the 13th International Conference on Information Fusion*, 2010.

[5] R. Yager, "A framework for reasoning with soft information," *Information Sciences*, vol. 180, no. 8, pp. 1390–1406, 2010.

[6] M. Kandefer and S. C. Shapiro, "An f-measure for context-based information retrieval," in *Commonsense 2009: Proceedings of the Ninth International Symposium on Logical Formalizations of Commonsense Reasoning* (G. Lakemeyer, L. Morgenstern, and M.-A. Williams, eds.), (Toronto, CA), pp. 79–84, The Fields Institute, Toronto, CA, 2009.

[7] S. L. Reed, "Texai." http://sourceforge.net/projects/texai/, 2010.

[8] J. R. Anderson, "Human associative memory," in *How Can the Human Mind Occur in the Physical Universe?*, pp. 91–134, NY, New York: Oxford University Press, 2007.

[9] M. B. Howes, *Human Memory*, ch. Long-term Memory – Ongoing Research, pp. 131–160. Sage Publications, Inc., 2007.

[10] D. Reisberg, *Cognition – 3rd Edition*, ch. Associative Theories of Long-Term Memory, pp. 251–287. W.W. Norton & Company, Inc., 2005.

[11] J. G. Romero, J. Garcia, M. A. Patricio, J. M. Molina, J. Llinas, S. C. Shapiro, M. Kandefer, and M. Prentice, "Strategies and techniques for use and exploitation of contextual information in high-level fusion architectures," in *Proceedings of the 13th International Conference on Information Fusion*, 2010.

[12] K. Sambhoos, J. Llinas, and E. Little, "Graphical methods for real-time fusion and estimation with soft message data," in *Proceedings of the 11th International Conference on Information Fusion*, pp. 1–8, 2008.

[13] A. M. Collins and E. F. Loftus, "A spreading activation theory of semantic processing.," *Psychological Review*, vol. 82, no. 6, pp. 407–428, 1975.

[14] F. Crestani, "Application of spreading activation techniques in information retrieval," *Artif. Intell. Rev.*, vol. 11, no. 6, pp. 453–482, 1997.

[15] A. S. Maida and S. C. Shapiro, "Intensional concepts in propositional semantic networks," *Cognitive Science*, vol. 6, no. 4, pp. 291–330, 1982.

[16] S. C. Shapiro, "Cables, paths and "subconscious" reasoning in propositional semantic networks," in *Principles of Semantic Networks* (J. F. Sowa, ed.), pp. 137–156, San Mateo, CA: Morgan Kaufmann, 1991.

[17] S. C. Shapiro and W. J. Rapaport, "SNePS considered as a fully intensional propositional semantic network," in *The Knowledge Frontier* (N. Cercone and G. McCalla, eds.), pp. 263–315, New York: Springer-Verlag, 1987.

[18] M. Prentice and S. C. Shapiro, "Usin propositional graphs for soft-information fusion," in *Proceedings of the 14th International Conference on Information Fusion*, 2011.

[19] A. N. Steinberg and G. Rogova, "Situation and context in data fusion and natural language understanding," 2008.

[20] S. C. Shapiro, "An introduction to SNePS 3," in *Conceptual Structures: Logical, Linguistic, and Computational Issues Lecture Notes in Artificial Intelligence 1867* (B. G. . G. W. Mineau, ed.), (Springer-Verlag, Berlin), pp. 510–524, 2000.

[21] J. R. Anderson, "Cognitive architecture," in *How Can the Human Mind Occur in the Physical Universe?*, pp. 3–43, NY, New York: Oxford University Press, 2007.

[22] D. Bothell, *ACT-R 6.0 Reference Manual*, 2010.

[23] D. B. Lenat, "Cyc: A large-scale investment in knowledge infrastructure," *Communications of the ACM*, vol. 38, no. 11, 1995.

[24] National Geospatial-Intelligence Agency, "NGA: GNS Home." http://earth-info.nga.mil/gns/html/index.html, 2010.

[25] C. van Rijsbergen, *Information Retrieval, Second Edition*. Butterworths, 1979.

[26] S. C. Shapiro, "Relevance Logic in Computer Science," in *Entailment* (A. R. Anderson, J. Nuel D. Belnap, and J. M. Dunn, eds.), vol. II, pp. pg. 553–563, Princeton, NJ: Princeton University Press, 1992.