# Information Fusion in a TV program Recommendation System

**Claire Laudy**
THALES Research & Technology
Palaiseau, France
claire.laudy@thalesgroup.com

**Jean-Gabriel Ganascia**
ACASA
Laboratoire d'Informatique de Paris 6
Paris, France
jean-gabriel.ganascia@lip6.fr

***Abstract*** ***– In a previous paper, we presented a conceptual framework for high-level information fusion. The approach relies on the use of the Conceptual Graphs model for both knowledge representation and information fusion. We introduced the use of fusion heuristics in the framework. This paper validates the viability of the approach through several experimentations. We use the fusion platform implementing the framework within a recommendation system for smart television. TV program descriptions coming from different sources of information are fused. To evaluate the quality of the fusion process, we compare the results to TV program descriptions recorded by the INA institute. We consider both the preciseness of the schedule and the proportion of programs correctly predicted. The experimentations emphasize the importance of the heuristics that we introduced in the conceptual framework.***

**Keywords:** Conceptual graphs, High-level information fusion, Ontology, Fusion strategies.

## 1 Introduction

The first step of a decision-making process is to gather the relevant pieces of information and to combine them in order to have a global representation of the external world. Such a process is difficult as information is distributed across various sources and on different media. For example, a terrorist attack may be reported through news, SMS's, videos, etc. This makes awkward the task of combining all the pieces of information into a coherent and accurate global view.

While most of the studies ([1], [2] and [3] for instance) concentrate on the fusion of low level data, our aim is to integrate heterogeneous information, thanks to the fusion of high-level information. The general approach that we propose is summarized on Figure 1. The observations coming from each source of information are first interpreted as information of a higher level of semantics. The high-level information items are then fused using domain knowledge, which has an even higher level of semantics.

The main issue is to integrate sources of information that are highly heterogeneous. The heterogeneousness of information concerns the formalism that is used and the level of semantics conveyed by each source. The media on which the information is communicated may also be different. Finally, the reliability of each source may differ.

In the Information Fusion community, recent papers report about how to use ontologies to store domain knowledge ([4]). Our work goes in the same direction: proposing techniques able to model knowledge.

On the modeling point of view, our approach relies on the use of conceptual graphs ([5]) for information and knowledge representation. Concerning the fusion process, we propose to take advantage of the operators defined on the conceptual graphs structures. We use one of these operators: the *maximal join*. We extend the standard maximal join, in order to take into account domain knowledge and user preferences. Therefore, we introduce the use of heuristics called fusion strategies inside the maximal join operation. Fusion strategies are rules encoding domain knowledge. They are used to extend the notion of compatibility between concepts.

This paper validates our approach through several experimentations. We used the fusion platform that we developed within a TV program recommendation system. The recommendation system computes TV program descriptions and decides whether a program should be recommended or not to a specific user. To this aim, we used the fusion platform in order to get TV program descriptions that are precise and reliable concerning both the scheduling times and the content description. We measured the quality of the results thanks to the comparison with reference information. We collected this reference information from low-level descriptions recorded by the INA institute.

Section 2 presents related works and reminds of the approach that we proposed in [6]. We describe the case study in section 3. Section 4 is dedicated to the validation of our proposition. Our experimentation protocol is described, as well as the tests that we conducted. We detail the results of the experimentations, emphasizing on the importance of the fusion strategies. We then conclude and present future work.

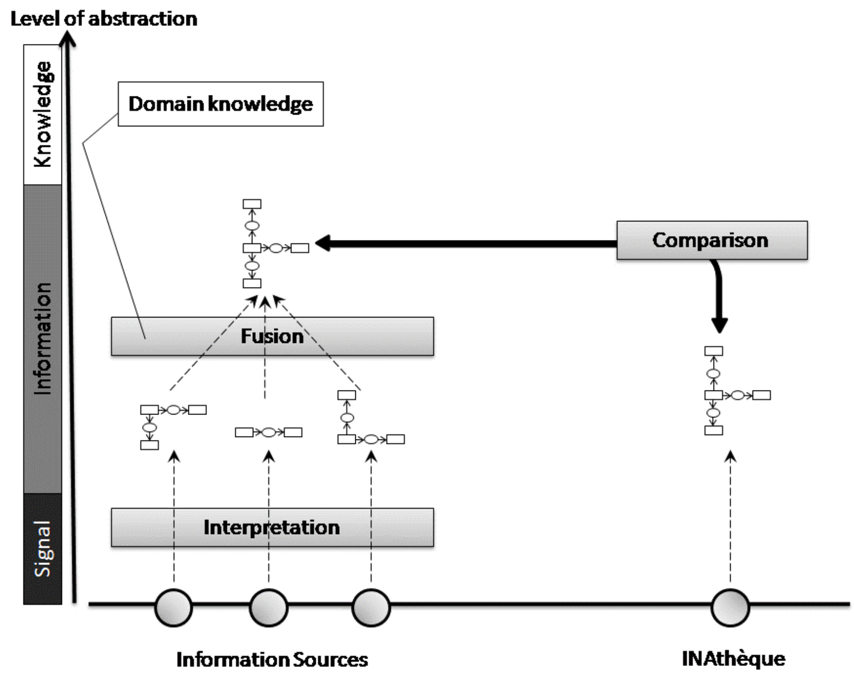


Figure 1: High-level Information Fusion approach

# 2 Context

## 2.1 Related Works

The general aim of our approach is to use the output of intelligent sensors as input observations for our system. For textual information, these intelligent sensors are systems able to analyze the meaning of the texts and store it as machine-readable information. As conceptual graphs were initially developed in order to analyze natural language, a lot of studies exist, aiming at transforming textual information items into conceptual graphs ([6], [8], [9]) or retrieving information from text using conceptual graphs queries ([10]). Considering other media, studies have been realized ([11], [12]) which aim at automatically analyzing images and videos and store the resulting descriptions as conceptual graphs. Finally, as stated in [13] and [14] conceptual graphs are widely used to formalize several domains of knowledge as different as biomedical risks or corporate modeling. Therefore, we use conceptual graphs as a common representation formalism for the storage of information coming from heterogeneous sources as well as for knowledge representation and domain modeling. Furthermore, we go beyond the usual use of conceptual graphs and take advantage of conceptual graph operators for information fusion.
Considering the fusion process, some studies aim at fusing information containing a high-level of semantics. The work reported in [15] concerns data integration for the semantic Web. Entities and relations are identified in data coming from texts, databases, and XML documents. The extracted entities are then compared in order to find the ones that refer to the same objects of the external world. Nevertheless, the focus of the work is on entities themselves. On the contrary, our aim is to identify interactions between objects, and thus relations between entities. In [16], information items are stored in a graph based representation. Pre-defined situations are described by multiple property regions and are mapped into points in a unit hypercube. The information items are mapped in the same representation and fused. The main difference with our proposition is on the understandability of the knowledge representation. We claim that human experts have to supervise the fusion process. Therefore, the knowledge representation must be understandable by domain experts who are not experts of representation formalisms.
The information fusion community is nevertheless more involved in studies aiming at fusing low-level data (for instance signal processing). The use of techniques and methods taken from natural language processing is a new field of interest in the fusion community (see [17] and [18] for instance). People look at how to use ontologies to model a domain. We claim that conceptual graphs are a good candidate for information fusion since the formalism contains the maximal join operator and the structures are easily understandable.

## 2.2 Using Conceptual Graphs for Information Fusion

In [6], we presented a framework for high-level information fusion that uses the conceptual graphs formalism. The Conceptual Graphs model is a formal model based on Existential Graphs [19] and Semantic Networks [20]. It was proposed by JF Sowa in [5]. The model is essentially composed of a support and the graphs themselves. The support is an ontology, which defines the vocabulary used in the graphs. A conceptual graph represents several concepts and the conceptual relations, which exist between them. Conceptual graphs are composed of entity and relation nodes. In the graphical representation, the entities are drawn as rectangles while the relations are ovals. The entities are either abstract or individual entities.

In our approach, the ontology and the canonical basis of the model are used to store the domain knowledge and the observations. In the ontology we describe all the entities of the external world and the relations that may be observed among them. Further knowledge about the domain, such as the situations that we expect to observe in the external world, is stored as abstract conceptual graphs in the canonical basis. The observations are acquired from several sensors and stored as instantiated conceptual graphs. The entities of the conceptual graphs representing the observations have referents or values.

Concerning the fusion process itself, it also relies on the conceptual graphs model. We use the maximal join operation defined by Sowa in order to fuse information. As shown in Figure 2, the maximal join operation allows to fuse two compatible subgraphs of two conceptual graphs. Graph G3 is the result of the fusion of G1 and G2 using the maximal join operation.

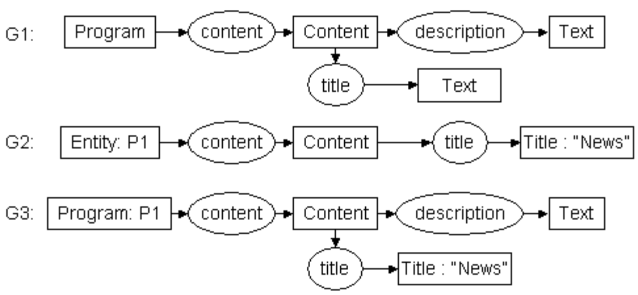


Figure 2: Example of Maximal Join operation

The maximal join operation copies all the information that is present in the initial graphs in the new one. Furthermore, it gives several results, which depict the different ways of combining the information, that is to say, the different fusion hypothesis.

Furthermore, as stated in [6], using the maximal join only is not sufficient in order to fuse information coming from real systems. Real data is noisy and knowledge about the domain is often needed in order to fuse two different but compatible values into a single one. Observations such as a person named “J. Smith” and a one named “Mr. John Smith” are not equals, but our background knowledge let us believe that the two observations rely to the same person. Therefore, we introduced the notion of fusion strategies. They are rules encoding domain knowledge and fusion heuristics. We use them to compute the fused value of two different observations of the same object. On the one hand, the fusion strategies extend the notion of compatibility that is used in the maximal join operation. According to some fusion strategy, two entities with two different values may be compatible and thus fusable. On the other hand, the strategies encompass functions that give the result of the fusion of two compatible values.

# 3 Case Study

The approach that we propose can be used on domains as different as medical diagnosis or business intelligence. A model of the domain has to be drawn *a priori* and stored as an ontology. We applied the approach within a TV program recommendation system. While the number of channels that one can access increases very fast, the aim of the system is to help the users in choosing the programs that they would enjoy seeing. Based on background information and the description of a new program, the system evaluates whether the new TV program is of interest to a specific user. The recommendation system can be coupled with a Personal Video Recording system, so that interesting programs are recorded even when the users are not looking at TV.

A new incoming TV program is evaluated on the basis of a description. The description must therefore be very detailed concerning the content of the program itself. It should also be as precise as possible concerning the broadcast times. Thereby the recording system can record the right slots of time.

An overall observation of the different sources of information that give TV program descriptions showed us that the TV program descriptions were either very poor considering the content of the program, or very unreliable concerning the broadcast times. The purpose of our study is to fuse the descriptions given by different sources. Our aim is first to obtain more complete and precise descriptions of the TV programs. The descriptions should contain both technical information and content description. Second, we want to obtain a better scheduling of the programs. The fused schedule should be as close as possible to reality.

To evaluate the preciseness of the schedule, we compared the fused descriptions to the ones stored by the INA (Institut National d’Audiovisuel [21]). The INA records all the descriptions and real broadcasting times of all the programs of the French TV and radio.

Within the recommendation system, we used two sources of information. The first one, called DVB stream, is the live stream of metadata associated with the video stream on the TNT (Télévision Numérique Terrestre). Figure 3 shows an example of the information available on the DVB stream.

The DVB stream gives descriptions of TV programs containing schedule and title information. It is very precise concerning the begin and end times of programs and delivers information about the technical characteristics of the audio and video streams. However, no description of the content of the program is given.

For each TV channel, it gives the descriptions of the currently playing program as well as the following one. The information on this source is constantly being updated. In particular, the scheduling times of the following programs are adjusted.

```xml
<?xml version="1.0" encoding="ISO-8859-1" ?>
<tv generator-info-name="TSReader">
[...]
  <channel id="1537-TF1">
    <display-name lang="en">TF1</display-name>
    <transport-stream-ID>6</transport-stream-ID>
    <signal-info>-0.0 MHz</signal-info>
  </channel>
  <programme start="20061127063959" stop="20061127064753" channel="1537-TF1">
    <title>Jt matin</title>
    <desc> |-0.0 MHz</desc>
  </programme>
  <programme start="20061127064754" stop="20061127083027" channel="1537-TF1">
    <title>TF ! JEUNESSE</title>
    <desc>Au sommaire  «Franklin». «Tabaluga». «Dora». «Bob l'éponge».|-0.0 MHz</desc>
  </programme>
[...]
</tv>
```

Figure 3: initial observation on TNT metadata

The second source of information is an online TV magazine. Figure 4 depicts an extract of the information delivered by this source. The descriptions contain information about the scheduling of the programs, their titles and the channels on which they are scheduled. They also contain details about the contents (summary of the program, category, actors, presenters etc). This source describes all the TV programs scheduled on all the TV channels during one week starting from the current day. The TV program descriptions may be updated once a day. As adjustments on the TV programs diffusion occur during the day, shifts often happen between scheduled times and real diffusion times.

```xml
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE tv SYSTEM "xmltv.dtd">
<tv source-info-url="http://telepoche.guidetele.com/" generator-info-name="XMLTV" >

  <channel id="C1.telepoche.com">
    <display-name>tf1</display-name>
    <icon src="http://static.guidetele.com/c_img/chaine/tf1.gif" />
  </channel>
[...]
 <programme start="20061127064500 +0100" stop="20061127083500 +0100"
showview="5755621" channel="tf1">
    <title>TF! Jeunesse</title>
    <desc lang="fr">Franklin. Tabaluga. Dora l'exploratrice(sous-titrage télétexte).</desc>
    <category lang="fr">emission jeunesse</category>
  </programme>
[...]
</tv>
```

Figure 4: initial observation on telepoche.fr

On the XML initial observations, we can see the information that we are going to fuse. For instance, the beginning time of the TV program appears inside the `<programme>` marker, as "start" attribute, the title is between the `<title>` markers, ... An XMLTV parser was developed within the case study, to automatically achieve the *Interpretation* step (see Figure 1), and transform the XML observations into conceptual graphs.

# 4 Validation

## 4.1 Implementation of a Fusion platform

In order to validate our approach, we implemented a fusion platform based on the use of conceptual graphs for knowledge representation and observation storage. As fusion algorithm, this platform uses the extension that we propose to the maximal join operation. The fusion platform was developed in JAVA and uses the AMINE platform ([22]).

AMINE is used as a service provider for conceptual graphs definitions and basic manipulations. Our platform implements a new version of the maximal join. It allows to call fusion strategies when joining two relations. It also gives all the results of the maximal join of two graphs. Indeed, each result is a fusion hypothesis and is useful for the global fusion and decision-making processes.

The fusion strategies are rules that were first implemented as JAVA classes. These classes are independent from the core platform itself. They are used as external functions.

## 4.2 Experimentations

As detailed before, the domain that we chose in order to validate our proposition concerns TV program descriptions. The aim is to obtain as much TV program descriptions as possible, concerning the TV programs scheduled on a TV channel, during one day. Furthermore, these descriptions should be as precise as possible with regards to the programs that were effectively played on the channel.

We realized the experimentations using sixteen TV channels. The sources of information for the program descriptions are the DVB stream of metadata and an on-line TV magazine.

In order to compare the results of the fusion to the programs that were really performed, we collected TV program descriptions from the INAthèque. The INA, Institut National de l'Audiovisuel, collects the descriptions of all the programs that have been broadcasted on the French TV and radio. The exact begin and end times of the different programs are recorded, as well as a brief description of the contents. To evaluate our results, we compare our fused descriptions to the INAthèque records. Thereby, we first know whether a fused program corresponds to the program that was really played. Second, we compare the times that were processed by fusion to the real diffusion times.

Our experimentation protocol is the following one. We request every 5 minutes the two sources of information to give us the next program scheduled on one channel. The two provided TV program descriptions are then fused using the fusion platform combined with one of the fusion strategies.

After fusion, we check that the description follows the general model for TV program descriptions stored in the canonical basis.

| | TF1 | France4 | France5 | BFM | Gulli | iTV | M6 | NRJ12 | NT1 |
|---|---|---|---|---|---|---|---|---|---|
| % prog found - no strategy | 0 | 15,52 | 59,7 | 7,15 | 1,13 | 37,67 | 0 | 0 | 0,4 |
| % prog found - strategy 1 | 0,18 | 18,53 | 73,47 | 7,15 | 1,13 | 42,71 | 0 | 0 | 9,66 |
| % prog found - strategy 2 | 51,49 | 92,24 | 90,34 | 42,29 | 80,23 | 73,97 | 32,5 | 36,16 | 77,05 |
| % prog found - strategy 3 | 64,23 | 92,24 | 80,46 | 30,08 | 71,35 | 82,99 | 47,71 | 37,5 | 81,88 |

Figure 5: Percentage of programs correctly fused and identified with different strategies

For instance, if the program has two different titles, the fusion failed because the strategies were not able to fuse the titles of the two observations. The observations most likely relied to two different programs and the resulting description is rejected.

The well-formed descriptions are then compared to the INA reference data. If the titles, subtitles, channels etc. are compatible, the fused program description is considered to be correctly found with regards to reality. If the description is either badly formed or any part of the description doesn't correspond to the reference data, we consider that the fusion failed. The program wasn't correctly found.

For correctly found programs descriptions, we then compare the computed begin and end times to the real ones. A smart TV program recorder can then use the fused descriptions. This system should automatically record selected TV programs. It means that the recorded programs shouldn't be ”cut” and the recording time should be as precise as possible in order to be useful to the user.

## 4.3 Fusion Strategies: the intelligence of the system

The quality of the fusion that we obtained using different strategies was measured. To this aim, we launched our experimentations using the fusion platform first combined with no strategy and then with three different ones. The first experiment -no fusion strategy- is equivalent to using the initial maximal join operator for information fusion.

The three strategies encode knowledge about the domain and are the following ones:

- **Strategy 1** extends dates compatibility. Two dates are compatible if the difference between the two is less than five minutes. If two dates are compatible but different, the fused date should be the earliest one if it is a ”begin date” and the latest one otherwise.
- **Strategy 2** extends dates and titles compatibility. The dates compatibility is the same as for strategy 1. Two titles are compatible if one of them is contained in the other one.
- **Strategy 3** extends dates and titles compatibility. The dates compatibility is the same as for strategy 1. Two titles are compatible if the length of the common substrings exceeds a threshold.

As first interpretation, we looked at the percentage of programs that were correctly found, according to the different strategies. By correctly found, we mean that the fused program description is compatible with the description stored in the INAthèque. Figure 5 shows the results that we obtained on a representative selection of TV channels.

As expected, we can see that the fusion of observations using the maximal join operation only is not sufficient. Only the descriptions with strictly identical values are fused. There is too much noise in real data for a fusion process that doesn't take into account some knowledge about the domain. Therefore, we applied the three previously cited fusion strategies. The more the compatibility constraints between two values are relaxed, the better the results are. It is equivalent to inject more and more domain knowledge in the fusion process. The issue then is to find the right amount of knowledge that has to be injected.

The low percentage of program correctly found on BFM is due to the specificity of this channel. BFM broadcasts almost only news and the titles of the different news programs are very similar. The fusion strategies are incongruous for this kind of channel. The problem of NRJ12 is different. NRJ12 broadcasts a lot of music. The DVB source describes every single video clip, whereas telepoche.fr describes longer programs containing several hours of video clips. Therefore, the observations are almost always incompatible.

## 4.4 Taking the context as a clue for fusion

A second interpretation of our results consisted in the observation of the time lag between the fused descriptions broadcast times and the reference ones. This gives us an idea whether the automatic recording system would be usable or not. Indeed, if the system cuts the programs, records too much advertisement or even records a non targeted program, users will not trust it and therefore not use it. We also measure the usefulness of this method to enrich data. Figure 6 and Figure 7 give examples of the results obtained on two different channels.

Each point represents a program and is located on the grid according to the difference in minutes between the fused begin and end times and the reference ones. The X-axis displays the time lag between begin times, while the Y-axis displays the time lag between the end times. Negative values indicate that the recording system would cut the program, whereas positive values indicate that it would record too much. On Figure 6 only three points are visible. Actually, only two programs were badly guessed and the point with coordinates (0,0) represents all the other programs. On Figure 7 we can see that all the programs are starting after the fused begin time. This is due to the fact that advertisement is scheduled at the beginning of the time slots dedicated to each TV program.

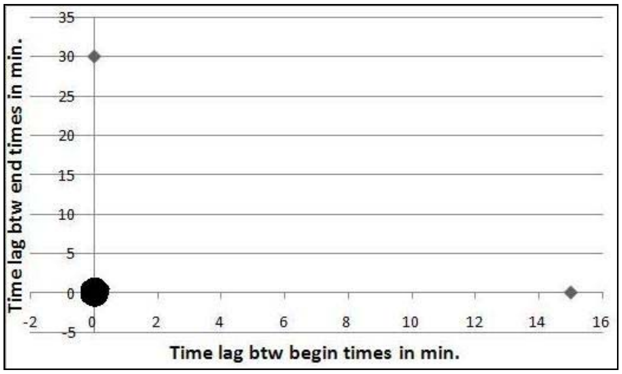


Figure 6: Time lag between fused and broadcasted times on France 4 channel

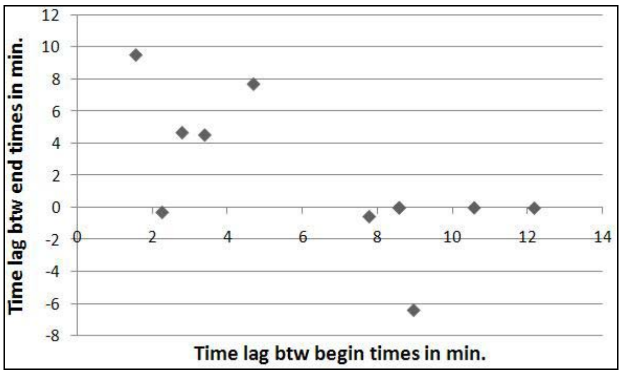


Figure 7: Time lag between fused and broadcasted times on TF1 channel

The different experimentations that we carried out showed that the quality of the fusion process is heterogeneous, according to several parameters. First of all, it depends on the channel on which the observations are done. Indeed, some channels broadcast the programs almost always at the scheduled time. The observations on both sources of information are identical (with regard to the times) and coherent with reality. In the meantime, most channels don't follow this rule. This can be due, for instance, to the fact that they broadcast a lot of live shows, have specific advertisement policies or are sensible of current events and may therefore change their programs at any time.

Other parameters on which the fusion results can be dependant, is the period of the day and the specificity of the channel. For non-popular channels (BFM…) and at periods of low audience (early morning), we observed a lot of errors in the programs given by the TV magazine.

To emphasize on the need to adapt the strategies to the context of the observations, we launched another experimentation. The aim was to show the difference of quality of the results between the use of a fusion strategy that doesn't take the context of observation into account, and the use of one taking that context into account. Therefore, we looked more precisely at the TF1 channel. During our first experimentations, we observed that the advertisements were generally scheduled at the beginning of the slots of times given by the TV magazine source. Therefore, we encoded a strategy that explicitly says that if the observations are made in the context of TF1 channel (i.e. the entity "Channel" of the conceptual graph has the value "TF1"), then the DVB source of information is more reliable concerning the begin times. The remaining of the strategy is unchanged, with regards to strategy 3.

Figure 8 shows the results that we obtained. It depicts the time lag between fused and real diffusion times for programs broadcasted on TF1. Compared to Figure 7, we can see that the quality of the fused beginning times is better. According to the types of programs and to the time of the day, an even smarter strategy could be used.

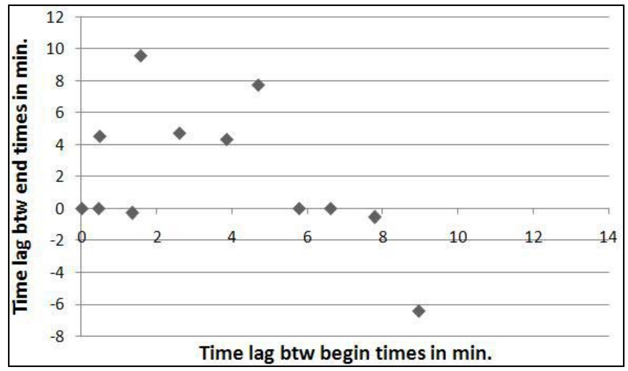


Figure 8: Time lag between fused and broadcasted times on TF1 channel – specific strategy on begin times

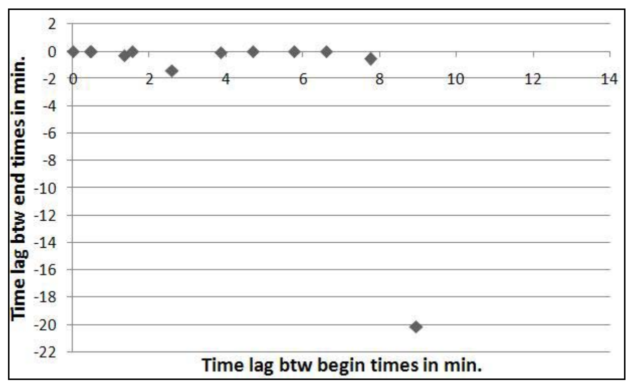


Figure 9: Time lag between fused and broadcasted time on TF1 channel – specific strategy on begin and end times

Figure 9 shows the results using a second context-aware strategy. As for the begin times, the end times are the ones given by the DVB source of information when TF1 is observed. As we can see, the results are almost always better. Only one program that was already badly guessed using the other strategies is worst. Two different episodes of a same series were fused.

## 4.5 Including Identification Strategies

As said before, the results of the fusion of TV programs that are scheduled on periods of low audience are very bad. Among other particularities, we observed that the TV magazine has "holes", especially for non-popular channels. During such periods, as next program to be broadcasted, the magazine source of information gives a program that will actually be broadcasted several hours after, whereas, the DVB gives the real next one.

| | TF1 | France4 | France5 | BFM | Gulli | iTV | M6 | NRJ12 | NT1 |
|---|---|---|---|---|---|---|---|---|---|
| % prog found no Identification str. | 50,78 | 77,16 | 76,2 | 54,5 | 67,95 | 70,06 | 37,14 | 35,77 | 58,41 |
| % prog found with Identification str. | 97,25 | 93,1 | 70,38 | 97,2 | 78,77 | 76,17 | 53,48 | 66,15 | 76,5 |

Figure 10: Percentage of programs correctly fused with and without identification strategy

The two descriptions are then incompatible and the resulting fused program is not well formed.

To overcome such problems, we introduced the use of another type of strategies: the identification strategies. Their role is to identify the observations that really correspond to the object that has to be observed. They encompass functions able to check the compatibility of two observations as well as functions able to choose the observation to return to the user in case of incompatibility.

Figure 10 shows the different percentages of program correctly found, first using no identification strategy, then using one based on title similarity and distance between begin and end times. When two descriptions correspond to two different programs, we select the program that is scheduled first (begin and end times are earlier). Through the results, we can see the usefulness of this new kind of strategy.

# 5 Conclusion

After reminding of the approach that was previously proposed for Information Fusion, this paper emphasizes on the experimentations that were conducted. The information fusion approach relies on the use of the conceptual graphs model for knowledge and observations representation. For the fusion process in itself, it relies on the use of an operation defined on conceptual graphs : the maximal join. The maximal join was extended with fusion strategies in order to introduce the use of domain knowledge in fusion. Fusion strategies are rules that allow to add a domain dependent notion to the fusion process.

Through this paper, we validate the viability of our approach. We used the approach within a recommendation system. We fused the TV program descriptions coming from two different sources in order to obtain more detailed and reliable ones.

The first results of our study are promising as we showed that the use of the maximal join operation is relevant for information fusion. The operator must nevertheless be enriched with domain knowledge in order to be useful on real data, which are noisy.

Through the different experimentations, we showed that the intelligent part of the system lies on the fusion strategies. Indeed, they are the place where domain knowledge and context is taken into account in order to be used as fusion heuristics. We also showed the usefulness of taking the context of observation into account during the fusion process.

As we said before, the strategies are the intelligent part of our fusion system. Therefore, they don't need to be only simple rules. They can encode complex functions, such as heuristics that were already used in other fusion studies and are designed for specific domains or specific kind of observations. Future work will concern the development of the identification and fusion strategies.

Concerning the identification strategies, they rely on similarity processing between both the graph structures, and the concepts on the other hand. Several studies such as [23] handle the problem of similarity between (conceptual) graphs structures. Similarity between concepts will rely on the use of measures such as the one proposed by [24]. The similarity processing between concepts will also take into account the similarity between the values of the concepts referents. We already looked at similarity measures for string referents, for the fusion of program titles for instance. We looked at studies addressing problems such as identifying duplicate records in databases. We used the Levenshtein edit distance [25]. Other smarter measures, such as learnable measures ([26]) can be used. The measure could be generic for any string comparison and parameterized thanks to a learning phase on a set of data specific to the application domain.

Concerning the fusion strategies the problem of uncertainty and reliability of the sources of information has to be considered. In our first experimentations, the reliability of each source is used inside the fusion strategies, but no level of uncertainty or reliability is attached to the fused results. We envision to integrate the use of belief functions ([27] and [28]) inside the fusion process to handle such problem.

In order to show the genericity of our approach, future work will be dedicated to the use of the fusion platform on other application domains. It will first be used in the context of an Intelligence system. We are currently studying a crisis management case study concerning the Ivory Coast events. Information items are extracted from several newspaper articles and fused in order to get an overview of the situation in the country at different times.

# References


[1] M. Hanif, U. Ali, *Optimized Visual and Thermal Image Fusion for Efficient Face Recognition*, 9th International Conference on Information Fusion, Florence, Italy, 2006.

[2] R. S. Blum, *Minimax robust image fusion using an estimation theory approach*, 8th International Conference on Information Fusion, Philadelphia, PA, July 2005.

[3] E. Dalberg, A. Lauberts, R. Lennartsson, M. Levonen, L. Persson, *Underwater target tracking by means of acoustic and electromagnetic data fusion*, 9th International Conference on Information Fusion, Florence, Italy, 2006.

[4] C. Matheus, M. Kokar and K. Baclawski. *A Core Ontology for Situation Awareness.* 6th International Conference on Information Fusion, Cairns, Queensland, Australia, 2003

[5] J. F. Sowa, *Conceptual Structures. Information Processing in Mind and Machine*, Addison-Wesley, Reading, MA, 1984

[6] C. Laudy, JG. Ganascia, C. Sedogbo, *High-level Fusion based on Conceptual Graphs*, 10th international Conference on Information Fusion, Quebec, Canada, 2007.

[7] P. Zweigenbaum, and J. Bouaud., *Construction d'une représentation sémantique en Graphes Conceptuels à partir d'une analyse LFG*, 4ème Conférence sur le Traitement Automatique des Langues Naturelles, Grenoble, France, 1997.

[8] J. Villaneau, J-Y. Antoine, and O. Ridoux, *LOGUS : un système formel de compréhension du français parlé spontané-présentation et évaluation*, 9ème Conférence sur le Traitement Automatique des Langues Naturelles, Nancy, France, 2002.

[9] M. Montes-y-Gomez, A. Gelbukh, A. Lopez-Lopez, *Text mining at detail level using conceptual graphs*, 10th international conference on conceptual structures, Borovets , Bulgaria, 2002.

[10] D. Genest. *Recherche d'information par transformation de graphes dans le modèle des graphes conceptuels*, Ingénierie des systèmes d'information, volume 7, numéro 1-2, pages 207-236. Hermès, 2002.

[11] P. Mulhem, and W. K. Leow, and Y. K. Lee, *Fuzzy Conceptual Graphs for Matching Images of Natural Scenes*, 7th International Joint Conference on Artificial Intelligence, Seattle, Washington, USA, 2001.

[12] M. Charhad, *Modèle de Documents Vidéo basés sur le Formalisme des Graphes Conceptuels pour l'Indexation et la Recherche par le Contenu Sémantique*, Thèse de L'université J. Fournier, Grenoble, 2005.

[13] F. Volot, and M. Joubert, and M. Fieschi, *Knowledge and Data Representationwith conceptual graphs for Biomedical Information Processing : a Review,* Methods Inf Med. , N°37 pp. 86-96, 1998.

[14] O. Gerbe, and B. Guay, and M. Perron, *Using Conceptual Graphs for Methods Metamodeling*, 4th International Conference on Conceptual Structures, Bondi Beach, Sydney, Australia, 1996

[15] J. Starz, B. Kettler, P. Haglich, J. Losco, G. Edwards, and M. Hoffman, *The Concept Object Web for Knowledge Management*, 4th international semantic web conference, Galway, Ireland, 2005

[16] J. T. Rickard, *Level 2/3 Fusion in Conceptual Spaces*, 9th International Conference on Information Fusion, Florence, Italy, 2006.

[17] F. Deloule, D. Beauchêne, P. Lambert, B. Ionescu, Data Fusion for the Management of Multimedia Documents, 10th international Conference on Information Fusion, Quebec, Canada, 2007.

[18] M. Gagnon, Ontology-based Integration of Data Sources, 10th international Conference on Information Fusion, Quebec, Canada, 2007.

[19] C. S. Pierce, *Collected papers of Chalres Sanders Pierce*, volume 2, C. Hartshone & P. Weiss (Eds.), Cambridge, MA, 1932.

[20] F. Lehman, *Semantic Network in Artificial Intelligence*, Pergamon Press, Oxford, UK, 1992.

[21] INAthèque: www.ina.fr/archives-tele-radio/universitaires/index.html

[22] AMINE Platform: amine-platform.sourceforge.net/

[23] S. Sorlin, P-A Champin, C. Solnon *Mesurer la similarité de graphes étiquetés* In proc. of 9èmes Journées Nationales sur la résolution pratique de problèmes NP-Complets, pp 325-339, 2003

[24] J. Zhong, H. Zhu, J. Li, Y. Yu. *Conceptual Graph matching for Semantic Search*, in Proc. Of 10th international Conference on Conceptual Structures, pp. 92-106, Borovets, Bulgaria, 2002.

[25] L. I. Levenshtein, *Binary codes capable of corecting deletions, insertions and reversals*. Soviet Physics Docklady, 10, 707-710.

[26] M. Bilenko and R. J. Mooney. *Adaptive duplicate detection using learnable string similarity measures*. 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, 2003.

[27] P. Smets. *Data fusion in the Transferable Belief Model*. In Proc. of 3rd International Conference on Information Fusion, Philadelphia, pp. 21-33, Paris, France, 2000.

[28] P. Vannoorenberghe. *Un état de l'art sur les fonctions de croyance appliquées au traitement de l'information*. Revue I3, Volume 3, N° 2; pp 9-45 (2003)