Interoperability of Information Fusion Systems

Mieczyslaw M. Kokar
Jakub J. Moskal
VIStology, Inc.
Framingham, MA, USA
mkokar@vistology.com
jmoskal@vistology.com

Information Fusion (IF) studies solutions for combining information from multiple sources into one representation that is more concise, sound, and informative than any source individually or a disjointed union of all sources. In this paper, we assess three issues that are currently hampering the interoperability of IF systems: 1) Ontology of IF, 2) Formal Theory of Information Integration and Fusion, and 3) Situational Awareness.

We show a generic information fusion (IF) system in Figure 1. It inputs two (or more) sets of information (Info 1, Info *n*) from the sources (Source 1, Source *n*), *represented* in some data structure, and inserts them (Integrate) into one data store (Info 1 ⊕ Info *n*). This is just a disjoint union of the two data sets. The next step (Compose) aligns the data items that represent *the same* information, followed by the integration and fusion part of the algorithm that combines those data items into new assertions about the world. This process may also involve assigning degrees of "belief" to each assertion. Figure 1 also shows the operation of *inference,* indicated by the Derive labels. The sizes of the rectangles representing inferred sets of facts indicate how much new information is inferred by the specific Derive engines. The intent here is to show that the amount of information after the operation of Compose is larger than K1, K*n*, and K1+K*n*, while Info is more concise than Info 1 ⊕ Info *n*.

Since IF involves multiple information sources, each represented in different schemes, there is a need to first align the schemes. Usually, IF is concerned with resolving the coordinates of an object that is detected by a sensor. IF normally starts with a concrete representation of all the variables that are relevant to the problem, and the data are collected as values of such variables. The variables, typically, are of the type Real, and the names of the variables are provided by the designer of an IF system. The mappings between two sets of real numbers are obvious: 5 = 5 and 3 ≠ 5. This is because the language used does not have any ambiguities. However, if the information is presented in a language that does not have types (like Real numbers) but rather uses linguistic terms, it is necessary to align them [1]. The reason for this is that nonmathematical languages do not follow the principle of unique name assumption (UNA), in which two items with the same identifier (ID) are the same and two items that have different IDs are not necessarily different (e.g., Tom ≠ Bob), might need to be explicitly asserted.

In addition, the open-world assumption (OWA), often used in logic-based languages, posits that an observer can never have complete knowledge and therefore cannot deem something false merely because it lacks evidence of its truth. In such a case, the agent can only say that the status of a statement is "unknown". An opposite to OWA is the closed-world assumption (CWA). In CWA, inherent in database systems, if a piece of information is absent from the database, it is automatically false. Consequently, fusion agents that accept UNA, OWA, or CWA may arrive at vastly different decisions.

These aspects result in a requirement for alignment of types and then composition of multiple information sets into one in a consistent way. Thus, not only do the items from two information sources (Info 1, Info *n*) need to be treated as one, but the implications of such associations also need to be analyzed. For instance, if one source has an object Alice, the other has an object Bob, and the alignment states that (Alice = Bob), then all instances of Alice and Bob in each of the two information sets need to be replaced with a term that is not used for something else. However, since both Alice and Bob may occur in different

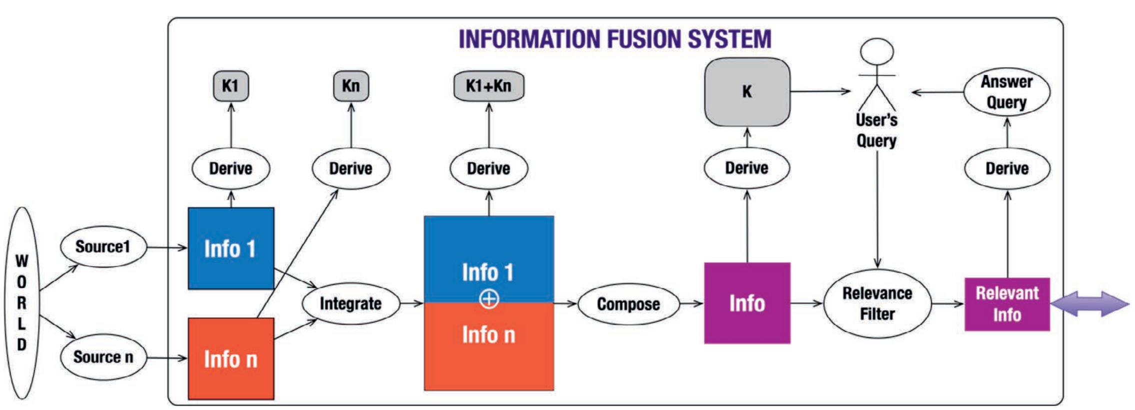

Figure 1
A conceptual view of an information fusion system.

relations with other objects, e.g., <Bob friend Tom> and <Alice enemy Tom>, this may cause a logical inconsistency, provided that the system knows that (friend ≠ enemy). Then, such an alignment is not admissible. However, if we want to align Bob with Tom and if the system knows that <Bob friend Alice> and <Tom friend Carl>, the system should infer that the object called Bob is friends with both Alice and Carl. In summary, alignment should be admissible if it does not imply inconsistencies, and alignment can result in the inference of relations between the aligned objects and other objects.

ONTOLOGY OF IF

If IF systems, like the one shown in Figure 1, were interoperable, they would be able to exchange information and knowledge shown in the figure (subject to the policy of the owner of the system)—information interoperability. The systems would need to use the language that they can understand, and the language should be standardized—a valuable goal for the IF society. Such a language is called *ontology*. Ontologies have been used in IF for a long time; see, e.g., [2], [3]. The use of ontologies has become quite popular. However, so far, we have not seen an ontology of IF that follows the ontology standards of the World Wide Web Consortium.

Such an ontology would need to provide for: (1) representation of object *types* (e.g., sensors, vehicles, and people), processes/functions, and their *instances*; (2) classifications of objects and processes, e.g., fusion functions; (3) representation of the Joint Directors of Laboratories (JDL) and levels; (4) representation of relationships between them; and (5) capability to specify IF systems and apply formal inference to *deduce* other facts that are only implicit in the ontology describing the systems. Ontologies then could be used to represent specific IF systems, and the representations could be used to infer *relations between the systems*. An example used in our paper was a *data fusion* system modeled as a subclass of a *decision fusion* system [4].

We are not aware of the existence of a generic ontology for IF. Although there are papers that refer to "information ontology" and "data fusion ontology", they do not satisfy all five of the above criteria. On the other hand, there are many papers that use specific ontologies inside of their IF systems.

THEORY OF INFORMATION INTEGRATION AND IF

The next level of interoperability would be the reuse of algorithms from one system in another. However, any algorithm interacts with many other algorithms within the system, so when we plug an algorithm into a system, it needs to be integrated with the whole system. To achieve this level of interoperability, IF needs methods of *algorithm composition* and *integration*. In our early searches for a theory of IF, we identified *category theory* as the most appropriate tool to be used as a most general model of integration of information; we introduced it to the IF community in [4]. We showed that algorithms cannot be combined by set-theoretic operations like union or product, but a more abstract category theory provides means for doing this—*morphisms, limits,* and *colimits*. We have used category theory in some of our work. A deeper theoretical investigation of fusion using category theory in general, and sheaves in particular, can be found in [5].

SITUATIONAL AWARENESS

We consider an IF system agent (or user) to be aware when they can answer queries about specific situations. Thus, a higher level of interoperability is to exchange information about situations among the IF systems and let the information receiver infer answers to the user queries (Figure 1). Within the JDL Data Fusion Model, which is widely used within the IF community, situational awareness is positioned at level 2 as a process and labeled Situation Assessment. The basis for this process is Endsley's work, e.g., [6]. In this conceptualization, "situation assessment" is understood as a process that can be measured and evaluated. If the process is efficient, then a high level of "situation awareness" is achieved. However, e.g., [3], [7] consider situations as objects, which have their own existence and thus can be described, and their descriptions can be exchanged, learned, and so on. The awareness of an agent of a specific situation is assessed then as the capability of answering queries *about the situations*. As shown in Figure 1, the process of Relevance Filter is part of an IF system. Its objective is to identify which parts of information and knowledge are relevant to a specific user's query. Relevant Info is then the description of the situation the user is inquiring about. It can be conveyed to another agent—either human or computer—for their use. In this scheme, Relevant Info is much smaller in size than all the information in the IF system; thus, it is less demanding on both the bandwidth of the communication channels and the user's cognitive load.

REFERENCES

1. Dapoigny, R., and Barlatier, P. Formal foundations for situation awareness based on dependent type theory. *Information Fusion*, Vol. 14, 1 (2013), 87–107.
2. Boury-Brisset, A. C. Ontology-based approach for information fusion. In *Proc. 6th Int. Conf. Information Fusion,* Vol. 1 (July 2003), 522–529.
3. Kokar, M. M., Matheus, C. J., and Baclawski, K. Ontology-based situation awareness. *Information Fusion,* Vol. 10, 1 (2009), 83–98.
4. Kokar, M. M., Tomasik, J. A., and Weyman, J. Formalizing classes of information fusion systems. *Information Fusion,* Vol. 5, 3 (2004), 189–202.
5. Robinson, M. Sheaves are the canonical data structure for sensor integration. *Information Fusion,* Vol. 36 (2017), 208–224.
6. Endsley, M. R. Toward a theory of situation awareness in dynamic systems. *Human Factors,* Vol. 37, 1 (1995), 32–64.
7. Lu, S., and Kokar, M. M. Inference-based information relevance reasoning method in situation assessment, 2023, under review.