TOWARD A GENERIC ARCHITECTURE FOR MULTISOURCE INFORMATION FUSION

John R. Josephson and B. Chandrasekaran
Ohio State University
Laboratory for Artificial Intelligence Research
Department of Computer & Information Science
Columbus, Ohio

Mark Carroll
Aetion Technologies LLC
Columbus, Ohio

ABSTRACT

This paper sketches a design for software for building, composing, and maintaining information fusion application software. It is intended to support a broad range of information fusion applications that vary with respect to types of information sources (imaging sensors, human reports, etc.) and with respect to operational demands (e.g., timeliness, reliability, availability of human attention). The design relies on abductive inference, compositional modeling, model-based prediction, anomaly-driven processing, and need-driven processing. Some main components of the design have previously been demonstrated in prototype systems that have shown good performance.

INTRODUCTION

According to Hall and Steinberg (2001, p 21.4), the limitations of current practices in multisensor data fusion include the following: complex signal propagation, dense target environments, background clutter, lack of training data, no real fusion of image and non-image data, lack of reliable methods for knowledge engineering, difficulty of modeling rapidly evolving situations, difficulty in controlling incommensurate sensors, and difficulty of linking human needs to sensor control. To this we might add: weakness in image fusion, and lack of systematic methods for plugging things together to provide continuous high-level awareness in ways that ensure that all-source systems are controllable, maintainable, and upgradeable.

This paper attempts to sketch how all these limits might be pushed back using model-based methods and by imitating some features of human perceptual and conceptual information processing. In this design, models contribute both to the modularity needed for effective engineering, and to the representations needed to capture relevant knowledge of targets, confusables, aggregates, sensors, and physics. Imitation of perception and cognition contributes principles for organizing and controlling the hierarchy of interpretive functions that are required to climb the levels of abstraction from raw sensor data, through object tracking, object identification, and situation assessment, to provide continuously updated representations to support threat assessment, recognition-primed response and intelligent planning.

SUMMARY OF THE BASIC IDEA

The basic idea can be described as "model-based information fusion" based on a tight integration of abductive inference (best-explanation reasoning) with modeling and simulation. Automated abductive inference will be used to create a changing, "best interpretation" representation of the situation from incoming data. Modeling and simulation will be used by the abductive-inference software for automatic generation of predictions from hypotheses, enabling the continual generation of predictions to support: hypothesis evaluation, sensor tasking, planning, and detection of anomalies that may be valuable indications of deception, modeling errors, or sensor failure. The combined system will use domain knowledge encoded as composable model fragments to provide a self-correcting representation of the situation based on current evidence from sensors and other forms of data.

Processing is organized into "layers" where the conclusion of one layer become the data to be explained at higher layers. Information generally flows bottom-up, as data items stimulate explanation-seeking processes. However, information also flows top-down driven by information-seeking goals such as those that arise for discrimination among rival hypotheses, refinement of hypotheses, and

Prepared through collaborative participation in the Advanced Decision Architectures Consortium sponsored by the U. S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-0009. The U. S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon.

response to human information needs. Lower-level conclusions may also be affected by expectations that arise from higher-level conclusions.

MODELING AND SIMULATION

The success of fusion of information from multiple sensors critically depends on the establishment of a common reference frame for the input data, usually achieved by geo-referencing of both the location of sensors and their fields of view. This sets up the geometric transformations that link the “viewer centered” coordinate systems of images and sensory fields with an “object centered” coordinate system in which surfaces and entities are co-referenced, and in terms of which disparate sources of information can be fused (Marr 1982). Spatial tracking enables an entity to be co-referenced over time, and so enables information fusion at the level of entities, rather than simply at the level of events in space and time. Entity-level information fusion facilitates entity recognition based on behavior as well as recognition based on entity-centered intrinsic characteristics, such as 3D shape, rather than only from single-perspective, sense-specific and typically 2D appearances. Once co-reference is established at the entity level, non-spatial information can be linked, such as information about identity, type, capabilities and intent.

Models are useful for recognition, for descriptions of complex situations, and for prediction of behavior. However, to achieve all of these functionalities with the same family of representations, it must be possible to compose situation models from parts that are adapted from elements in repositories of prestored knowledge, and it must be possible to simulate the composed models to derive detailed consequences and expectations for specific situations. Thus, there is a need for technology for compositional modeling and simulation, with representations that are sufficiently expressive and domain independent to represent entities, features and parts of entities, entity types, sensors, sensor characteristics, signals, and the causal processes that modulate signal characteristics and mediate the effects of entities on sensors.

CFML (Compositional Functional Modeling Language), under active development at Ohio State and Aetion Technologies, is nascent technology for modeling and simulation designed to have the desired properties of representational expressiveness and compositionality, while supporting multiple types of “smart” simulation. CFML is inspired in part by Stanford's Compositional Modeling Language (CML), described in Falkenhainer, *et. al.* (1994). CFML extends and adapts CML in a number of ways, the most important conceptual advance being the addition of abstraction relations. CFML also incorporates capabilities for run-time equation solving that are more complete and robust than CML’s.

CFML, like CML before it, represents knowledge in the form of “causal model fragments,” which represent entities and relations. Causal constraints are represented as equations, which may apply conditionally, depending on discrete states. Equations are solved at run time by heuristic methods to derive dependent values to answer specific questions. Under human or machine control, entities can be: described at multiple resolutions, composed using defined relations, embedded in described environments, given initial conditions and constraints, and simulated to answer specific questions. Complex interactions can be modeled between large numbers of diverse entities that can change their states with corresponding changes in constraints on behavior. An arrangement of entities and relations describes a situation. For each situation, predictions can be automatically derived for plausibility testing, anticipation, monitoring, and discovery of anomalies. Large numbers of situation hypotheses that potentially explain current data may be considered in parallel.

CFML’s simulation engine uses both symbolic and numerical methods to derive the values of dependent variables from values of independent variables, where the order of dependency may be determined at run-time. In terms of knowledge, the function of the simulator may be described as using causal knowledge to deduce predictions for hypothetical situations, which may be complex. Although CFML was designed for representing engineered artifacts, it is very general, and it is a good candidate technology to represent the situations and entities about which information needs to be fused for a given set of tasks.

In CFML, an alternative situated in an environment is said to be *deployed*, and specifying it requires specifying the alternative, the environment (with its properties) and relations between the alternative and the environment. This is similar to composing an alternative. Once an alternative is deployed, specific questions may be asked. Questions can be answered using static analysis, or using time-based simulation, where time steps are advanced until there are no unanswered questions, or a time limit is exceeded. For example, we can ask about how a particular type of thing might appear in our data. CFML is “object oriented” and similar in this way to many other languages for simulation. However, the orderly methods for composing objects from parts, and situations from objects and environments, are unique to CFML. Variables may take values that are continuous or discrete, numeric or symbolic. This is similar to other advanced languages for simulation and modeling.

ABDUCTIVE INFERENCE

Inference to the best explanation - **abduction** - is a distinctive and recognizable pattern of evidential reasoning. It is ubiquitous at or near the surface of typical arguments offered in science, intelligence analysis, and ordinary life, and may be considered part of commonsense logic. An abductive argument is open to attack in characteristic ways, and may be defended in characteristic ways by supporting arguments. Abductive arguments are fallible, but there are only a small number of ways in which they can go wrong. This analysis provides a framework for justification, criticism, and dialog concerning the evaluation of evidence that can be supported in software, and used to capture the evidential basis for many types of assertions. While abductive arguments are not the only argument forms that are important in reasoning about situational hypotheses, they are central, and pervasive, and their specific strengths and weaknesses are critical determiners of the quality of information fusion software.

An abductive argument has a pattern that can be described as follows:

D is a collection of data (facts, observations, givens).
Hypothesis H explains D (would, if true, explain D).
<u>No other hypothesis explains D as well as H does</u>.
Therefore, H is *probably* correct.

Note that the conclusion is justified, not simply as a possible explanation, but as the *best explanation in contrast with alternatives*. The strength of the conclusion H, the force of the *probably* in the conclusion statement, reasonably depends on the following considerations:

- how decisively the leading hypothesis surpasses the alternatives,
- how good this hypothesis is by itself, independently of considering the alternatives,
- how thorough was the search for alternative explanations.

Besides the judgment of likelihood, willingness to accept the conclusion also reasonably depends on pragmatic considerations, including:

- how strong the need is to come to a conclusion at all, especially considering the possibility of gathering further evidence before deciding,
- the costs of being wrong and the rewards of being right.

Humans intuitively recognize these five considerations. Moreover, humans sometimes actually come up to the standards set by these considerations in their actual reasoning. That is, whether or not the arguments are made explicit, people commonly come to conclusions that can be defended using strong arguments addressed to these considerations, with arguments for the distinct superiority of the leading hypothesis, for the deficiencies of alternative explanations, and so on. Thus, this serves both as a descriptive account of human reasoning and argumentation, and also as a prescriptive account, describing characteristics of logically strong abductions, and suggesting that logically strong abductions are indeed *good reasoning*, and *tend toward truth*.

The failure and success of predictions enters naturally into the evaluation of hypotheses. A hypothesis that leads to false or highly inaccurate predictions is poor by itself, and should not be accepted, even if it appears to be the best explanation when considering all the available data. Failures in predictive power count as evidence against a hypothesis and so tend to improve the chances of other hypotheses emerging as best. Failures in predictive power may improve the margin of decisiveness by which the best explanation surpasses the failing alternatives. Thus, abductions are capable of turning negative evidence against some hypotheses, into positive evidence for alternative hypotheses. Besides the success of predictions, several other factors may enter in to the evaluation of a hypothesis, either in isolation or in contrast with rivals. These factors include: internal consistency, plausibility (including precedent and consistency with background knowledge), likelihood (prior, after local-match evaluation, and all things considered), simplicity, explanatory power, specificity, and productive promise. One problem with narrow probabilistic approaches to abduction is a tendency to represent all reasoning over these considerations as numerical computations in the one dimension of likelihood.

It is important to distinguish abduction, considered statically, as a pattern of argumentation or justification, from the dynamic inferential processes whereby explanatory hypotheses are generated and evaluated, and from the processes of constructing arguments, including abductive arguments. Many forms of information processing that are commonly used in automatic recognition, classification, and tracking algorithms can be seen as methods for performing abductive inferences, and although justificatory arguments are usually not explicitly generated, this could be done in principle to support debugging and after-action review.

It is also important to recognize that best explanations are almost always composite hypotheses or "theories" comprised of many parts. Coherent "scenes" and "situations" are best represented as composite hypotheses, as are theories in science and the meanings of sentences.

ABDUCTION MACHINES

Reasoning strategies for composing, criticizing, and revising hypotheses, and for justifying conclusions, can be investigated scientifically by implementing such strategies in software and testing their performance. Josephson and colleagues have been conducting this kind of research for several years, as described in the book *Abductive Inference* (Josephson and Josephson,. 1994, 96) and in many published papers. *Abductive Inference* describes six generations of "abduction machines" whose combined capabilities include: forming hypotheses by instantiation and by composition; forming maximally plausible composite hypotheses; removing explanatorily superfluous parts; determining the appropriate relative confidence for hypothesis parts; making use of elementary hypotheses with varying levels of specificity organized into a taxonomic hierarchy; justifying conclusions; decomposing larger explanation problems into smaller ones; ensuring a consistent composite explanation; opportunistic control; assembling low-specificity hypotheses, then controlling their refinement; using explicitly represented causal and structural knowledge to derive hypothesis-matching knowledge, to derive explanatory relationships, and to assemble causally integrated composite hypotheses; parallel processing; minimizing backtracking by forming the most confident partial conclusions first, then descending gradually from higher to lower confidence partial conclusions, all the way to intelligent guessing; forming confident partial explanations with ambiguous data left unexplained; using incompatibility relationships to reduce ambiguity and to extend partial explanations to more complete ones; handling hard and soft incompatibility and implication relationships; coordinating multiple sites of hypothesis formation while integrating bottom-up and top-down flow of information; gracefully handling noise and novelty; and forming a composite hypothesis that follows a causal chain from effects to causes.

The use of automated abductive inference of this power and sophistication to assist information fusion has never been done. The potential benefits are very great, especially given that previous work (e.g., *Abductive Inference, p. 223 ff.*) shows that abductive inference can use explanatory relationships to dramatically reduce uncertainty compared with recognition-based methods alone. The prospects for automating abductive reasoning are especially good for closed-world tasks such as recognition of well-defined behaviors of well-described objects. The difficulties are substantial and the technical risk is significantly greater in attempting to automate abductive reasoning in "open-world" domains that require reasoning about human agents, and behavior, and purposes. An important fallback from the goal of full automation is partial automation, which will still substantially reduce the burden on human processing and improve productivity, while leaving some tasks and subtasks, especially the most creative, for humans to perform. The prospects for effective automation of information fusion, especially in open-world domains, will be substantially improved if hypotheses are represented in such a way that composite hypotheses, especially large and complicated composite hypotheses, may be automatically simulated to derive predictions. We suggest that CFML, described in the last section, is a good start at what is needed.

PERCEPTION AS INFORMATION FUSION; PERCEPTION AS LAYERED ABDUCTION

Perception typically uses multiple data channels: for instance, it appears that primate vision has separate information processing channels for color, texture and shading, each supplying data to different layers of interpretation. Auditory perception combines information from the two ears, and also separates the signals into frequency bands. Functionally, combining information from different senses is similar to combining information from different channels within a single sense. In principle, information derived from one channel, sensor, or field report can be used to assist in deriving information from another, the combination being more powerful than just summing the contributions. For example, binocular hearing can detect and localize a sound source in 3-D space, attracting the eyes, which may provide n abundance of descriptive information about the object and a better spatial localization, and initiate change in the direction the sensor platform (head), which orients it toward the object, thus enhancing further pickup of information, both by the eyes and the ears. A key to integrating this information is thus reference to a common reference frame, or at a minimum the ability to relate different reports to one another.

Abductive Inference (Chapter 10) proposed a **layered-abduction computational model of perception** that purports to unify bottom-up and top-down processing in a single logical and information-processing framework. Two main hypotheses were proposed: 1. Perception has the same basic logic as theory formation in science (& diagnosis, & Sherlock Holmes, etc.); 2. Because they have similar information-processing demands, perception and theory formation use similar computational strategies (however, perception is more "compiled" and less flexible). In general, there are three levels of explanation for information-processing phenomena (adapted from Marr, 1982):

1. Content theory of the information-processing task.
   a. Input/Output - what is computed - the goal.
   b. How it is computed.
      - information content – knowledge
      - computational architecture
      - logic of the strategy
2. Realization theory - how it is computed: algorithms, data structures, networks.
3. Implementation theory - how it is computed: underlying mechanism - software, hardware, wetware, virtual machine.

In these terms, the layered-abduction level-1 content theory of perception can be given as:

Input: reports from data acquisition agents
Output: representation of reality (aspects of interest)
How computed: layered abductive inference

- Information content, knowledge: taxonomy of entity types, & relationships, characteristic entity features
- Architecture: layered, conclusions of one layer become data to be explained at the next layer
- Logic of the strategy: abduction - the conclusion at each layer is a well-justified best explanation based on the best available evidence at that layer

Abductive Inference proposed a specific account of the core computational strategies of generic perception. In brief, at each layer, a best-explanation composite hypothesis is formed of the data provided by lower layers, using information from higher layers to assist with identification and disambiguation. Processing occurs concurrently, in a distributed fashion. Local inferences seed islands of relative certainty which grow in three directions:

- to higher layers, becoming data to be accounted for;
- within layers, using hypothesis-hypothesis interactions of incompatibility, positive association, and logical implication;
- to lower layers, becoming expectations that guide the generation and evaluation of lower-level hypotheses.

Anomaly detection and targeted data collection

The proposed abduction machine for information fusion will use data about the world as a basis for combining hypotheses into a best composite hypothesis about elements of the current reality. Parts of this working model of the current situation (**WMCS**) composite hypothesis may be more confident than others, with confidences ranging from the equivalent of "more likely than not," through "highly probable," to "beyond reasonable doubt." "Could be" hypotheses can be made available, and also a facility for "intelligent guessing" that uses could-be hypotheses to extend the WMCS beyond the range of more-likely-than-not hypotheses to include hypotheses that are "more likely than anything else." In general, this WMCS hypothesis will be continually maintained and revised, rather than generated afresh. Elements of it will be presented to users at appropriate levels of detail to support their understanding and decision making.

Model-based predictions from the WMCS hypothesis prepare the WMCS hypothesis for revision. Predictions are compared with incoming data, and discrepancies are detected. Minor discrepancies cause small parametric adjustments in the updated WMCS hypothesis, whereas larger discrepancies are treated as anomalies that need to be qualitatively explained, and the WMCS hypothesis revised accordingly. A new entity may have entered the field of attention, or the type, or intent, or capabilities of an entity may have been mistaken. On the other hand, a data source may have become unreliable.

Sometimes the absence of an information report is highly significant. It may indicate the absence of its source, or suggest that its source is not providing coverage of a critical part of the space. It may show that an enemy has unexpectedly become capable of stealth, or that things are not what they seem, for example, that an entity is a decoy. The subtle but critical evidential importance of anomalously missing information is illustrated by the Sherlock Holmes story of *Silver Blaze* (Arthur Conan Doyle, 1892), where the key to the solution was that *the dog did not bark*, indicating that the thief must have been known to the watch dog.

> "Is there any point to which you would wish to draw my attention?"
> "To the curious incident of the dog in the night-time."
> "The dog did nothing in the night-time."
> "That was the curious incident," remarked Sherlock Holmes.

Thus, by using simulation to follow models of the world to their consequences, an abduction machine can reason about the deviation of data from expectations, including the absence of data that should be there if a data source is reliable and correctly tasked.

Dynamic abductive reasoning sometimes requires additional data to help in forming, assessing and refining hypotheses. These are specific functional needs that arise at certain points in the processing. At these points, precompiled knowledge of distinctive features can be used by an abduction machine to suggest what data should be gathered to fill the information needs of discriminating among rival hypotheses for regions of ambiguous data, or to fill in important missing details, for instance, to gather additional information about a particular event or entity.

BIBLIOGRAPHY

Falkenhainer, Farquhar, Bobrow, Fikes, Forbus, Gruber, Iwasaki, & Kuipers, *CML: A Compositional Modeling Language,* technical report, Knowledge Systems Laboratory, September, 1994. Available from http://ksl-web.stanford.edu/KSL_Abstracts/KSL-94-16.html.

Hall, D. L., & Steinberg, A. N. (2001). Dirty Secrets in Multisensor Data Fusion. In Hall, D. L., & Llinas, J. (Eds.), *Handbook of Multisensor Data Fusion*. (pp. 21.1-21.12), CRC press.

Josephson, J. R., & Josephson, S. G. (Eds.). (1994, 1996). *Abductive Inference: Computation, Philosophy, Technology*. New York: Cambridge University Press. The first chapter is available online at: http://www.cis.ohio-state.edu/~jj/pubs/AbdInfCh1.pdf

Josephson, J. R. (2001). On the Proof Dynamics of Inference to the Best Explanation. *Cardozo Law Review*, 22(5-6), 1621-1643.

Josephson, J. R. (2000). Smart Inductive Generalizations are Abductions. In P. A. Flach & A. C. Kakas (Eds.), *Abduction and Induction.* (pp. 31-44). Dordrecht/Boston/London/: Kluwer Academic Publishers.

Josephson, J. R. (1998). Abduction-Prediction Model of Scientific Discovery Reflected in a Prototype System for Model-Based Diagnosis. *Philosophica*, 61(1), 9-17.

Josephson, J. R., Chandrasekaran, B., Carroll, M., Iyer, N., Wasacz, B., Rizzoni, G., Li, Q., & Erb, D. A. (1998). An Architecture for Exploring Large Design Spaces. In *National Conference of the American Association for Artificial Intelligence*, (pp. 143-150). Madison, Wisc.

Josephson, J. R. (1995). Formation of Well-Justified Composite Explanatory Hypotheses Despite the Threat of Combinatorial Explosion. In R. Valdez-Perez (Ed.), *AAAI Spring Symposium on Systematic Methods of Scientific Discovery*, (pp. 133-138). Stanford, CA: AAAI. AAAI Technical Report

Marr, D. (1982). *Vision*. San Francisco: W. H. Freeman.

Murphy, R. R. (1996). Biological and Cognitive Foundations of Intelligent Sensor Fusion. *IEEE Transactions on Systems, Man and Cybernetics*, *26(1)*,

Schenk, T., 2000. Object recognition in digital photogrammetry. *Photogrammetric Record*, 16(95), 743--762.

Schenk, T., 1999. *Digital Photogrammetry*, vol. 1, TerraScience (www.TerraScience.net), 428 p.

Schenk, T., 1995. A layered abduction model for building recognition. In *Automatic extraction of man-made objects from aerial and space images*, Grün, Kübler, Agouris (Eds.), Birkhäuser Verlag, 117--124.

DISCLAIMER