Situations and Contexts

***Abstract*—**The semantics of context are examined, considering concepts of relevance, situations, and relationships. We define a situation as a set of relationships and a context as a situation that a) provides expectations for constituent entity states or b) is deemed relevant to the solution of an inference or response problem. The use of context variables in inferencing is examined. Predictive models as used in inferencing are construed as estimates of state distributions. The uses of context in inferencing can be differentiated into categories of target and information source characterization methods, appropriate to different assumptions concerning the quality of available prior models and observational data.

Alan N. Steinberg
Independent Consultant
Lake Ridge, VA, USA
alaneilsteinbeg@gmail.com

CONTEXT

Human understanding is infused with a robust sensitivity to context. Our sense impressions are informed by a myriad of mitigating and extenuating circumstances that enrich our experience and deepen our understanding. Consciously or unconsciously we search for contextual clues and use them to resolve ambiguous or puzzling situations.

Many shortcomings in attempts at artificial intelligence—in machine vision, robotics, natural language, understanding, and information fusion—as well as in understanding human behavior artificial intelligence can be attributed to deficient appreciation of context. In this article we explore methods (a) to define and represent context, (b) to determine contexts as relevant to particular uses, and (c) to incorporate contextual information in reasoning and decision making.

A recent survey of context-related literature reveals a diversity of definitions of *context* [1]. In some usages, a context is considered to be a *situation* of some relevance (as "the bombing can be understood in the context of the Middle East Crisis"). In others, it is an *element* of such situations (as "the enhanced security measures make sense in the context of the recent bombing"). In yet other uses, a context is *information* about a situation or even a *source* of such information.

"A context is a situation that provides information that can be used either a) to condition expectations or b) to improve the understanding of a given inference or planning/control problem".

We have suggested the following definition as conducive to understanding and using contexts: A *context* is a *situation* that provides information that can be used either a) to condition expectations or b) to improve the understanding of a given inference or planning/control problem [2]. These two ways in which a situation can be used as context derive from a formulation by Gong [3], contrasting notions of *context-of* (C-O) and *context-for* (C-F). A situation can be C-O or C-F, depending on how it is used in reasoning. C-O-driven reasoning starts with a perceived situation to derive expectations about constituent entities, relationships, and activities. In contrast, C-F-driven reasoning starts with a particular problem—which might be an inferencing problem (what's happening?) or a control problem (what's to be done?)—and seeks to discover additional information that can resolve uncertainties in the problem solution [4], [5].

RELATIONS, RELATIONSHIPS, AND SITUATIONS

If contexts are situations that can be used in inferencing, we need to understand what situations are and how to reason about them. As in [6], [7], we follow Devlin [8] in defining situations in terms of relationships.

Let us differentiate the concepts of *relation* and *relationship*. We shall use "relation" to designate an abstraction, such as *marriage, ownership, hatred,* or *selling.* "Relationship", on the other hand, is used to designate an instantiation of a relation anchored within a situational context: Antony's marriage with Cleopatra, Othello's marriage with Desdemona, or Cleopatra's marriage with Othello. As the latter examples indicate, such contexts are not necessarily factual either in the real world or in a particular assumed fictional context.

Reasoning about attributes, relations, relationships and situations is facilitated if these concepts are "reified", i.e., treated as entities in the working ontology [8]. Explicitly defined, a *relation* is a mapping from *n*-tuples of entities ($n \geq 1$) to a relational state *r*:

$$R^{(n)}: X_1 \times \ldots \times X_n \rightarrow \mathcal{R}. \tag{1}$$

A *relationship* is an instantiation of a relation; i.e., an ordered $(n+1)$-tuple $<r^{(n)},x_1,\ldots,x_n>$ such that *r* applies to a sequence of arguments $<x_1,\ldots,x_n>$. Attributes of individual entities are conveniently treated as unary relations, instantiations thereof as unary relationships.

The mapping from n + 1-tuples to relationships can be many-one, because the same entities $\langle x_1,\ldots,x_n\rangle$ may be related multiply by the same relation: Two companies may simultaneously have two contracts with one another. The type of relationship may be the same, at some level of abstraction, e.g., land-use contract—but the relationships—the individual contracts—are distinct. Therefore, we will want to differentiate variables and their values from instantiations thereof [9].

By reifying relations and relationships, we allow higher-order variables: those that range over predicates of other variables. These can be employed using the cross-order predicate of *application*, which we represent by parentheses. Thus, an expression in the familiar form "$x(i)$" is read as "attribute x applies to the individual i", e.g., Isaac is blind. Similarly, "$x(i,j)$" says that relation x applies to the individuals i,j, e.g., that Isaac is the father of Jacob. We can also distinguish between predicative variables X and particular values x thereof: as in "Isaac's height is 180 cm": "$H(i) = h$, $h = 180$ cm". This allows us such expressions as "$g(H)$", as in "*height* is a unary relation (i.e., an attribute)" or "*taller than* is a binary, transitive, nonreflexive relation".

We define a *situation* as a *set of relationships*. A *concrete situation s* is a set of fully anchored relationships $\{r|r$ obtains (i.e., holds true) in $s\}$. We should allow for imprecisely defined situations: The relationships that comprise a given situation may constitute a fuzzy set [6]. It will be convenient to conflate unit set and member to say that relationships are situations. What is more, because single-place relations are allowed, a single event (as in "the enhanced security measures are appropriate in the context of the recent bombing") is also a situation, consistent with our suggested definition of context.

Like relationships, a situation s may be real, or it might be conditional, hypothetical, fictitious, or otherwise counterfactual in some encompassing situation $t \supseteq s$ (t may be the universe at large). For example, the killing of Polonius is a situation that occurs in the context of Shakespeare's *Hamlet* but not in the world at large. The play *Hamlet* and various performances thereof, its plot, script, and various copies exist in the world at large. Its characters, situations, and events do not. The particular concerns of some agent (e.g., a person or an automated inference system) determine which situations are under consideration as contexts for those concerns.[1]

We can abbreviate $r \in s$, where r is a relationship and s is a situation as the conditional relationship $(r|s)$ read "r obtains in situation s". This is related to Devlin's use of the implicature notation $s \models \sigma$ for an infon σ [8].

THE USE OF CONTEXT IN INFERENCE

An inference problem q can be stated in terms of a utility function on the values of a problem-specific set of variables: $\omega_q : X \rightarrow \Omega^*$, where X is either a problem variable or a vector of problem variables. A *context for an inference problem* is a situation that is selected (by some agent) for use in understanding or solving the problem. We take this usage of context for an inference problem as an application of C-F, as used by Gong [3] and by us in [1], [2], [4], [5]: s is a context for resolving X, where X is a set of random variables. This can be contrasted with cases of C-O as in "in the context of today's economic news, it is likely that the Euro will strengthen". Such constructions have the form "in the context of S, $f(x)$", where P is a proposition (say, "the Euro will strengthen") and $f(.)$ is a modal expression, such as P is true or the probability that $P = p$ or it is likely that P or is impossible that P.

"The relevance of contextual information can be stated in terms of the contribution of such information in resolving values of problem variables…A context variable is a variable that an agent selects to evaluate or refine an estimate of one or more problem variables".

The relevance of contextual information can be stated in terms of the contribution of such information in resolving values of problem variables. We discuss this in the following. Let us consider how contexts can be used in evaluating problem variables to meet objectives. A general distinction can be drawn between *refinement* and *inference* of values of variables. In many data fusion problems, multiple measurements of a given variable are averaged or filtered to refine the estimate of that variable, exploiting independence in the measurement-to-measurement noise. Bayesian and Dempster-Shafer classifiers are examples of refinement (filtering) processes.

Often, however, the problem variables to be estimated are not themselves measured or are not measured with sufficient accuracy or confidence to meet users' needs. In such cases, the values of problem variables may be inferred totally or partially on the basis of other variables. Such inference assumes a model of the dependencies between measured variables and problem variables. Inference methods include, for example, structural equations, Bayesian belief networks, and neural networks.[2]

We may distinguish, then, between explicit problem variables and ancillary variables used in inference. We call the latter

[1] This we take to be the intent of Devlin's informal definition for situation as "a structured part of reality that is discriminated by some agent" [8, p. 31, paraphrased]. However, the agent should not be part of the definition: much like the noise of a tree falling in the forest, a situation can exist without being noticed or cared about. It is on this basis that we distinguish contexts from other situations.

[2] Kalman filters and related tracking filters are typically hybrid refinement or inference processes such that problem variables are those constituting a target's physical state (e.g., its kinematic state), but filtering occurs not in state space but in measurement space: received and predicted measurements are filtered to infer target states (by means of motion and measurement models), from which additional (e.g., future) measurements are predicted.

*context variable*s. A context variable is a variable that an agent selects to evaluate or refine an estimate of one or more problem variables. Accordingly, we can define a *problem context* as a situation, comprising a set of entities and their relationships involving context variables and problem variables. Situations are selected as problem contexts for their presumed usefulness in solving the particular problem.

When a situation is used as a C-O, context variables are situational variables (ranging over relationships and sets of relationships); when used as a C-F, context variables are variables that are other than a given set of problem variables. By this definition, one problem variable can serve as a context variable for evaluating another problem variable. For example, an aircraft's observed speed may be used as a context for resolving its type, and, conversely, its estimated type can be used for resolving its speed (e.g., in bearings-only target tracking) [10].

One way of defining relevance is statistical relevance as introduced by Salmon [11], whereby that the relevance of a value *y* of candidate context variable *Y* in determining a specific value *x* of problem variable *X* is

$$Rel\left(y,x\right)=\frac{p\left(x\mid y\right)}{p\left(x\mid\neg y\right)}. \tag{2}$$

Statistical relevance in a context *s* is, of course, given as

$$Rel\left(y,x\mid s\right)=\frac{p\left(x\mid s,y\right)}{p\left(x\mid s,\neg y\right)}. \tag{3}$$

The utility to a given inference problem *q* of evaluating a variable *Y* for the purpose of evaluating a problem variable *X* in the context of a situation *s* is

$$\omega_q\left(Y;X\mid s\right)=\omega_q(x\mid s)\int_X\int_Y Rel\left(y,x\mid s\right)dy\,dx. \tag{4}$$

For discrete-valued variables, integration can be replaced by summation. *X* or *Y* in this formulation can be an individual variable or a vector of variables. For example, the set of variables *Y* = {day of week, weather conditions, location} can provide a useful context for resolving joint states of interest in the set of problem variables *X* = {traffic conditions, location}.

CONTEXT-SENSITIVE STATE ESTIMATION

State estimation functions differ broadly according to the types of state variables to be estimated. It can be convenient to distinguish entity states of interest according to the "levels" described in various versions of the Joint Directors of Laboratories data fusion model. Levels of data fusion and resource management processes map into a categorization of entity state variables that a data fusion system is tasked to estimate or that a resource management system is tasked to control. Examples of such problem variables are given in Table 1. The third and fourth columns distinguish continuous-valued and discrete-valued variables at each level. The fifth and sixth columns, respectively, relate these to our particular rendition of data fusion and resource management levels [6], [7], [12].

"...it is preferable to distinguish inference problems on the basis of type of entity state variables rather than by type of entity. Depending on one's interests, many an entity can be considered alternatively as an individual (characterized in terms of level 1 variables) or as a relational structure (level 2 variables) or as a dynamic process (level 3). If it is a resource of the inference system itself, the same entity could be evaluated in terms of level 4 variables".

As we argue in [2], [7], it is preferable to distinguish inference problems on the basis of type of entity state variables rather than by type of entity. Depending on one's interests, many an entity can be considered alternatively as an individual (characterized in terms of level 1 variables) or as a relational structure (level 2 variables) or as a dynamic process (level 3). If it is a resource of the inference system itself, the same entity could be evaluated in terms of level 4 variables.

Both C-O and C-F can play essential roles at any fusion level, but they are especially important in higher level fusion, in which variables of interest include relation, relationship, and situation variables that are not directly observable but must be inferred. Although a C-F is useful in evaluating specific attributive and relational states, a C-O provides a means for understanding expectations for and implications of such states. Generally, the larger context in which a problem is considered, the more fully will it be understood by being conditioned on a larger number of mutually independent context variables.

Level 1 fusion is concerned with *attributive* states; that is, with values of 1-place state variables, such as target location, type, or attributive parameters. In level 2 fusion, both attributive and *relational* states are pertinent, i.e., values of *n*-place state variables, $n \geq 1$. Belief networks can be used to propagate information among entities, relations, and the relationships in which they participate. Given our reification of relation and relationships, we can depict a level 2 hypothesis after the pattern of Figure 1. This figure is in the form of a factor graph, in which variables are represented as circles and functions on these variables are represented as squares [13]. Examples of such functions are causal conditions or conditional probabilities, but they can represent any relationship among variables. In our application, the functions are instantiations: individual entities, relationships, and situations, etc.

Likelihoods and state estimates can be propagated among the nodes of a level 2 hypothesis. Each node combines the ef-

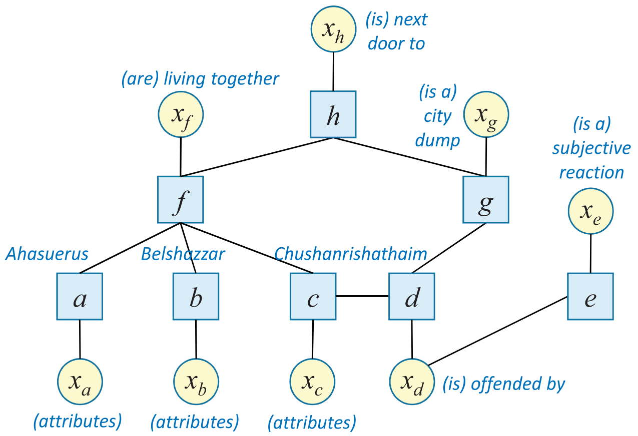

Figure 1
Factor graph representation of a level 2 hypothesis.

fects of evidence from its immediate neighbors and distributes its own evidence to them, ensuring, however, that information is not circulated back to an originating node. That is to say, a level 2 fusion process creates and updates nodes i in a level 2 hypothesis H, on the basis of data either

a) from an update to i based on associating one or more source reports with that node or

b) from an update to a node j that is an immediately neighbor to i in H.

In the former case, updating is attributive, analogous to updating a level 1 hypothesis. In the latter case, updating is relational, involving estimation of the relationships that occur between j and i in H and thereby refining the estimate of the state of i, x_i.

In the example shown in Figure 1, nodes *a, b,* and *c* postulate individual entities that participate in a relationship $f = <x_f,a,b,c>$ in which x_f is a relation. This relationship in turn, participates in another relationship $h = <x_h,f,g>$, in which g is yet another relationship. For example, x, y, and z could be the people Ahasuerus, Belshazzar, and Chushanrishathaim, and the relation x_f might be *living together* so that f is, roughly, a household or some subset thereof. Note that *living together* is an example of a relation of indefinite order, i.e., $x_f^{(n>1)}$. The relation x_h could be *next door to*; x_g the attribute (1-place relation) *city dump*; and x_d might be *is offended by*, applied to Chushanrishathaim in the relationship d. Additionally, x_e is a second-order attribute—perhaps *subjective reaction*—operating on the first-order relation x_d in the instantiation e.

The situation $s_1 = \{c, d, f, g, h\}$ might be the context for (C-F) evaluating Chushanrishathaim's attitude toward his present living situation. A different context could include his two housemates: $s_2 = s_1 \cup \{a, b\}$. Another might be the broader situation represented in the figure: $s_3 = s_2 \cup \{h, e\}$.

A belief propagation algorithm will determine the belief concerning the state of an entity (or, more precisely, of the vector of state variables associated with that entity) in terms of

Table 1

Entity State, Data Fusion, and Resource Management Levels[a]

Level	Entity Class	Example: Continuous State Variables	Example: Discrete State Variables	Data Fusion (Inference) Level	Resource Management Level
0	Patterns, e.g., features or signals	Temporal/spatial/spectral extent, amplitude, and shape/modulations	Signal/feature class, type, attributes	Signal/feature assessment	Signal/feature management
1	Individuals, e.g., physical objects or events	Location, velocity, size, weight, event time	Object class, type, identity, activity, or attributes	Individual entity assessment	Individual resource management
2	Structures, e.g., relationships and situations	Distance, force/energy/ information transfer	Class, type, identity, or attributes of relations, slots, arguments, situations	Situation assessment	Resource relationship management (coordination)
3	Processes, e.g., courses of action, scenarios, and outcomes	State utility, duration, transition conditions	State transitions; class, type, identity, attributes of processes, scenarios, or impacts	Scenario/ outcome assessment	Mission objective management
4	System resources	All of the above, applied to system resources	All of the above, applied to system resources	System assessment	System management

[a] [2], [7].

- "local" evidence $\varphi_i(x_i)$, i.e., information about a particular state variable and
- evidence $\psi_{i,j}(x_i,x_j)$ concerning the entity from other situation elements used as context.

If beliefs are expressed as probabilities, the joint probability distribution of the set of state variables $\{x_1,...,x_N\}$ corresponding to the N nodes in such a graph is

$$p\left(\left\{x_1,...,x_N\right\}\right)=\frac{1}{N}\prod_{(ij)}\psi_{ij}\left(x_i,x_j\right)\prod_i\varphi_i\left(x_i\right). \tag{5}$$

The function $\psi_{i,j}(x_i,x_j)$ is an undirected compatibility function—say, Pearson product moment correlation—as a generalization from the directed conditional probability $p(x_i|x_j)$ [14].

Evidence is propagated as "messages" passed to node i from nodes j in its immediate neighborhood $N(i)$ in the graph of relationships in the relevant situation:

$$b_i\left(x_i\right)=k\varphi_i\left(x_i\right)\prod_{j\in N(i)}m_{ji}\left(x_i\right). \tag{6}$$

Messages are updated recursively through the graph as

$$m_{ij}\left(x_j\right)\leftarrow\sum_{x_i}\varphi_i\left(x_i\right)\psi_{ij}\left(x_i,x_j\right)\prod_{k\in N(i)\setminus j}m_{ki}\left(x_i\right). \tag{7}$$

The evaluation over $k \in N(i)\setminus j$ in the last term of (7) indicates that data is to be passed from all immediate neighbors of i other than j itself. It is shown in [14] that such restriction on message passing maintains consistency and convergence in any singly connected (i.e., nonlooping) graph.

We can expand (7) by marginalizing over instantiated relation variables:

$$m_{ij}(x_j)\leftarrow$$
$$\sum_{x_i}\varphi_i(x_i)\sum_R p\left[x_i,x_j\mid R\left(x_i,x_j\right)\right]p\left[R\left(x_i,x_j\right)\right]\prod_{k\in N(i)\setminus j}m_{ki}\left(x_i\right). \tag{8}$$

This marginalization, of course, assumes discrete-valued relations. It is often practicable to partition continuous-valued attributes and relations into discrete bins for belief network propagation.

Because relations, attributes, and entities that are arguments of these can participate in multiple situations and relationships, the graph of a situation hypothesis can be multiply connected. Methods have been developed that provide exact or approximate joint probability distributions in a wide variety of graph topologies. These include Pearl's clustering algorithm [15], junction tree algorithms [16], the Shafer-Shenoy separator algorithm [17], and the generalized belief propagation formulation of Yedida et al. [14].

REASONING ACROSS FUSION LEVELS

As seen in Table 1, reasoning about relationships and situations has been considered the province of level 2 data fusion. Level 1 data fusion is concerned with estimation of states of entities considered as individuals. In contrast, fusion levels 2 and 3 are concerned with estimation of entities considered as aggregates: as relationships or situations and courses of action or scenarios, respectively [2], [7], [12], [18].

Situation assessment (level 2 data fusion)—whether implemented by people, automatic processes, or some combination thereof—involves inferences of the following types:

- inferring the presence and the states of entities on the basis of relationships in which they participate;
- inferring relationships on the basis of entity states and/or other relationships;
- recognizing and characterizing observed situations.

Whereas level 2 fusion concerns the estimation of observed states, level 3 fusion (concerns states that are projected; e.g., predicted future states [2], [7]. The temporal evolution of a situation, involving courses of action, interactions, and outcomes, constitutes a scenario [2], [18].

Level 2 and 3 inferences have direct analogy to those at level 1. Situation recognition is a problem akin to target recognition. Situation/scenario tracking is akin to target tracking [6], [18]. Characterizing situations is generally a matter of assessing the states of situation constituents and their interrelationships. The familiar Bayesian pattern for context-sensitive inferencing within fusion level 1 (L1 → L1) is given by

$$p\left[F(x)\mid G(x),s\right]=\frac{p\left[G(x)\mid F(x),s\right]p\left[F(x)\mid s\right]}{p\left[G(x)\mid s\right]},$$
$$p\left[G(x)\mid s\right]=\int_H p\left[G(x)\mid F(x),s\right]p\left[H(x)\mid s\right], \tag{9}$$

e.g., estimation of the probability of a single target state $F(x)$ from associated measurements $G(x)$ or prediction of state $F(x)$ from prior state $G(x)$ in situation s.

This can be generalized as

$$p\left[F^{(m)}\left(x_1,\ldots,x_m\right)\mid G^{(n)}\left(y_1,\ldots,y_n\right),s\right]=$$
$$\frac{p\left[G^{(n)}\left(y_1,\ldots,y_n\right)|F^{(m)}\left(x_1,\ldots,x_m\right),s\right]p\left[F^{(m)}\left(x_1,\ldots,x_m\right)\mid s\right]}{p\left[G^{(n)}\left(y_1,\ldots,y_n\right)\mid s\right]}; \tag{10}$$

with application to various inference patterns within and between fusion levels by selection of relation orders m and n (Table 2).

The reification of relations and relationships allows us to relate one to another in or out of context. In this way, attributes of relations and relationships can be inherited; e.g., in an L2 → L2 inference (with some uncertainty) from *x is providing information to y* to *x is cooperating with y.* A situation state cannot only imply but can be implied by the states and relationships of constituent entities so that situational inferences can be given in the form of Boolean combinations of expressions, such as

$$\exists r \exists x_1,\ldots,\exists x_n \left[r = R^{(n)}\left(x_1,\ldots,x_n\right) \& F_1\left(x_1\right) \& \ldots \& F_n\left(x_n\right) \& \{r,x_1,\ldots,x_n\} \subseteq s \right] \Rightarrow G(s). \tag{11}$$

CONTEXT IN MODEL ASSESSMENT AND MANAGEMENT

Data fusion relies, in one way or another, on predictive models of information sources and of entities of interest (targets at all appropriate state estimation levels). In military applications, target models are generally expected to be provided by intelligence processes. Operational intelligence, in a process called intelligence preparation of the battlefield, provides values for context-sensitive prior probabilities $p(x|s)$. Technical intelligence provides target characterizations that combined with source characterizations allow evaluation of measurement likelihoods $p(Z|x, g)$ and probabilities of detection $p_D(x|s)$ for target states x, information source state g, and measurement sets Z. Fusion systems use information source measurement models that combine with target descriptions to provide values for $p(Z|x, g)$ and $p_D(x|g)$ and with contextual information, e.g., target densities, and background clutter levels, to provide values for false alarm rates $p_{FA}(g, s)$ for given contexts s.

In many, perhaps most, current information exploitation systems, source models are the responsibility of source developers. However, there are many applications where valuable information is available from sources whose design and operating characteristics are unknown to the information exploitation system or its developers [19]. The wealth of information available online and from traditional open-source media provides enormous opportunities for diverse information exploitation applications, as do novel sensors and sensor platforms (drones, crowdsourcing, etc.) but require some means for quality control, as discussed in [1], [19].

We have reported on the development of a prototype information exploitation system that assesses and modifies the target and source models it uses at various fusion and management levels as a means of exploiting nontraditional information sources [20].

Model assessment is a level 4 data fusion process, performing all the classical data fusion functions:

- *data preparation:* aligning in feature space, structure, and confidence the data used in inferring models;
- *data association*: establishing the range of phenomena to be used in determining and validating the model; and
- *state estimation*: estimating the distribution and dependencies of characteristics and behavior of modeled entities or entity classes.

Model assessment differs in one major respect from level 0–3 data fusion processes and, indeed, from other level 4 fusion processes. The business of these other data fusion processes is the estimation of states of particular entities in the world; i.e., of *instantiations* of entity classes. In contrast, model assessment is concerned with the inference of *possible* states or, more precisely, the inference of the *distribution* of states possible for a given entity or class of entities. Model assessment performs estimation and prediction just as in other types of level 0–4 data fusion but with the difference that now the estimation and prediction are of the characteristics and behaviors of *distributions* of level 0–4 entities or of classes of such entities.

"Model assessment performs estimation and prediction just as in other types of level 0–4 data fusion but with the difference that now the estimation and prediction are of the characteristics and behaviors of distributions of level 0–4 entities or of classes of such entities".

Model assessment processing can take the form of induction from instantiated states to the distribution of possible states. It also can involve explanation of observed phenomena by subsumption to higher-level models. As an example, a radar performance model will gain in predictive and explanatory power to the extent that it is subsumed to electromagnetic physics and to which the latter is subsumed to unified quantum and relativistic physics. Source model management can involve setting parameters to compensate for estimated sensor biases (sensor regis-

Table 2

Relation Orders for Intra- and Interlevel Inferencing

Inference Type	*m*	*n*	Application
L1 → L1	1	1	Inferring states of an individual from states of the same or another individual
L1 → L2/3	1	>1	Inferring relationships from individual states
L2/3 → L1	>1	1	Inferring individual states from relationships
L2/3 → L2/3	>1	>1	Inferring relationships from other relationships

tration and calibration) and adjusting sensor accuracy models to reflect estimated error statistics. Target model management can involve modifying predictive models of target and situation classes in response to updated estimates of the characteristics and behaviors of such entities.

CATEGORIES OF TARGET CHARACTERIZATION PROBLEMS

There are numerous applications in which we cannot count on having high-fidelity models of target attributes or behaviors. For example, an adept, agile adversary, such as encountered in nonconventional warfare will not provide us with large samples of regular patterns of behavior for use in training statistical models. Such a problem is very different than conventional target recognition or tracking problems that can be addressed by model-based methods.

Waltz [21] has proposed a categorization of inference problems. We adapt this scheme in [2], [7] to distinguish inference methods by the way they use observational data and predictive models, as summarized in Table 3:

CATEGORY 0 (MODEL-BASED RECOGNITION)

This category encompasses methods used in traditional target recognition systems, relying on high-confidence models of target characteristics and behaviors. Prediction can involve deductive and inductive methods, whereby target entities and activities are recognized by matching observations to those predicted by models, possibly conditioned by the context of such factors as information source characteristics, viewing geometry, observation media, and background.

CATEGORY 1 (ANOMALY-BASED DETECTION)

It can happen that background (or normal) activities are better characterized than target activity. By matching observations with prior models of background activities, anomalous phenomena are detected as an indication of possible activities of interest.

Both categories 0 and 1 assume the availability of observational data and of prior models that have been validated in one way or another: In category 0, these are models of target entities or activities; in category 1 these are models of normal or background activities. Recognition and prediction (deductive and inductive) methods can be used in processing model data to derive expected observations for use in the matching process.

In contrast, category 2 and 3 methods are used to overcome deficiencies in prior models or in observable data, respectively. In category 2, new models are composed adaptively to explain observed data. In category 3, activities of interest might not be observable, rather their prior feasibility is determined on the basis of contextual information.

CATEGORY 2 (HYPOTHESIS-BASED EXPLANATION)

The process in this category is one of abductive reasoning: building and testing models to best explain available data. Such a process is applicable to situations in which there is insufficient prior analytic understanding or training data to develop predictive models. An analyst or an automated process constructs a situation or scenario hypothesis in an attempt to account for observed data. As in the classical scientific method, the hypothesis is evaluated to predict further observables that could either confirm or refute the hypothesis. By acquiring such data as available, explanatory, predictive models of the observed situation or scenario are selected, refined, or rejected.

"The uses of context in inferencing can be differentiated into categories of target and information source characterization methods, appropriate to different assumptions concerning the quality of available prior models and observational data".

CATEGORY 3 (CONTEXT-BASED FEASIBILITY)

These methods do not rely on direct observational data, rather, contextual cues are used to determine the feasibility of broad classes of activities: domain constraints on adversary capability developments, strategic planning, etc. Such methods are the only ones available when activities of interest are unlikely to be detectable or discriminable at all.[3]

[3] It might be useful to add yet another category (category -1?) to encompass estimation refinement via filtering or smoothing in the absence of a model; e.g., without model-driven filter gains.

Table 3

Categories of Inference Problems and Methods

Category	Approach	Assumed Prior Models	Observational Data	Inference Method
0	Model-based recognition	Targets	Yes	Deduction and induction
1	Anomaly-based detection	Backgrounds	Yes	Deduction and induction
2	Hypothesis-based explanation	Situation context and components	Yes	Abduction
3	Context-based feasibility	Targets and backgrounds	No	Deduction and induction

We can further subdivide category 0 to distinguish cases in which target models are "given" from those in which target models are derived by statistical learning:

- **0a** in which the actual target state (at whatever state estimation level) is known absolutely, e.g., under controlled test conditions;
- **0b** in which predictive models of targets and their behaviors are obtained explicitly from design documentation or are derivable analytically from first principles;
- **0c** in which predictive models are estimated from training data. In this problem category, distinct from category 0b, models are estimated inductively, for which all the apparatus of data fusion is applicable.

CATEGORIES OF SOURCE CHARACTERIZATION PROBLEMS

Inferencing problems involve the exploitation of information from sources whose performance may be well or poorly characterized. Categories of source characterization problems can be defined in terms of the availability of predictive models of source performance. This categorization is analogous to that for target state inferencing, as both reflect methods for acquiring knowledge concerning problem variables: level 4 variables in the source characterization case; level 0–3 variables in the target characterization case. As with the categories of target characterization problems, the categories of source characterization problems differ in their dependence on contextual information:

- **S0a:** in which the actual source performance is known absolutely, e.g., undercontrolled test conditions;
- **S0b:** in which predictive models of source performance are obtained explicitly from design documentation or are derived analytically from available information concerning the source's feature space and inference methods. Pertinent information of this sort may be reported by the source in real time, or it might be obtainable from source design documentation or from more general models of the source class (e.g., an analytic receiver model);
- **S0c:** in which predictive models are developed from training data, using estimates of source, target and situation states, together with historical performance data referenced to ground truth, i.e., historical measures of reporting errors in known conditions. We can further distinguish subcategories of S0c, based on the quality of available ground truth, as defined in terms of the above target characterization categories:
    - **S0c/0a:** source performance is estimated on the basis of observed entity states that are known absolutely, as in ideal test conditions;
    - **S0c/0b:** source performance is estimated on the basis of observed entity states that are well modeled;
    - **S0c/0b:** source performance is estimated on the basis of observed entity states that are inferred statistically;
    - **S0c/1:** source performance is estimated on the basis of observed entity states that are inferred from contextual anomalies;
    - **S0c/2:** source performance is estimated on the basis of observed entity states that are inferred by explanation of observable data;
    - **S0c/3:** source performance is estimated on the basis of observed entity states that are inferred by explanation of contextual data.
- **S1:** in which the performance of the given source is derived by comparison of its product with that from other sources. This category may be further refined by distinguishing these other sources according to their source characterization categories and by distinguishing degrees of independence among the sources (e.g., whether they measure or report commensurate variables);
- **S2:** in which predictive models of source performance are constructed abductively to explain the observed behavior of the source. Such a method is used when no reliable information is available concerning the source, but source performance must be inferred from target state estimates as reported by the source and compared with available ground truth (e.g., in in a test environment). Examples of reported state estimates include expectation and covariance matrices for continuous state variables (such as location or kinematics) and probability vectors across discrete state variables (e.g., target class within an exhaustive disjoint taxonomy);
- **S3:** in which performance of an information source must be inferred on the basis of context, i.e., from circumstantial evidence. Such methods can be necessary when reporting from the given source is so sparse and variable that it is not feasible to develop a predictive model of the given source. This can occur with obscure Websites, graffiti, and such "unsourced" information. Useful contextual information might include features of the source reporting medium and style, the known or assumed reporting conditions, and correlated reporting from other available sources. In some cases, it might be feasible to stimulate the source to observe its differential behavior under known conditions.

SUMMARY

We have attempted a careful definition of terms pertinent to discussion of situations and contexts. A context is treated as a situation that provides expectations for constituent entity states (C-O) or that is deemed relevant to the solution of an inference or response problem (C-F). Context exploitation involves a) predicting the value of contextual information to meet information needs; b) selecting information types and sources expected to provide information useful in meeting those needs; c) determining the relevance and quality of acquired information; and d) applying selected information to a given problem. Predictive models as used in inferencing are construed as estimates of state distributions. The uses of context in inferencing can be differentiated into categories of target and information source characterization methods, appropriate to different assumptions concerning the quality of available prior models and observational data.

REFERENCES

1. Rogova, G. L., and Steinberg, A. N. Formalization of "context" for information fusion. In *Context-Enhanced Information Fusion*, L. Snidaro, J. Garcia, J. Llinas, E. Blasch, eds. Zurich, Switzerland: Springer, forthcoming.
2. Steinberg, A. N., and Rogova, G. L. System-level use of contextual information. In *Context-Enhanced Information Fusion*, L. Snidaro, J. Garcia, J. Llinas, E. Blasch, eds. Zurich, Switzerland: Springer, forthcoming.
3. Gong, L. Contextual modeling and applications. In *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, Waikoloa, HI, 2005.
4. Steinberg, A. N., and Rogova, G. L. Situation and context in data fusion and natural language understanding. In *Proceedings of the 11th International Conference on Information Fusion*, Cologne, Germany, 2008.
5. Steinberg, A. N. Context-sensitive data fusion using structural equation modeling. In *Proceedings of the 12th International Conference on Information Fusion*, Seattle, WA, 2009.
6. Steinberg, A. N. Foundations of situation and threat assessment. In *Handbook of Multisensor Data Fusion*, M. E. Liggins, D. L. Hall, and J. Llinas, eds. Boca Raton, FL: CRC Press, 2009, pp. 437–502.
7. Steinberg, A. N. and Snidaro, L. Levels? In *Proceedings of the 18th International Conference on Information Fusion*, Washington, DC, 2015.
8. Devlin, K. *Logic and Information*. Cambridge, United Kingdom: Cambridge University Press, 1991.
9. Shapiro, S. C., and Rapaport, W. J. The SNePS family. *Computers & Mathematics with Applications* Vol. 23, 2–5 (Jan.–Mar. 1992), 243–275.
10. Angelova, D., and Mihaylova, L. Sequential Monte Carlo algorithms for joint target tracking and classification using kinematic radar information. In *Proceedings of the 7th International Conference on Information Fusion,* Stockholm, Sweden, 2004.
11. Salmon, W. C. *Statistical Explanation and Statistical Relevance*. Univ. of Pittsburgh Press, 1971.
12. Steinberg, A. N., and Bowman, C. L. Revisions to the JDL data fusion model. In *Handbook of Multisensor Data Fusion*, M. E. Liggins, D. L. Hall, and J. Llinas, eds. Boca Raton, FL: CRC Press, 2009, pp. 45–68.
13. Kschischang, F. R., Frey, B. J., and Loeliger, H. A. Factor graphs and the sum-product algorithm. *IEEE Transactions on Information Theory*, Vol. *47*, (2001), 498–519.
14. Yedida, Y. S., Freeman, W. T., and Weiss, Y. Understanding belief propagation and its generalization. In *Exploring AI in the New Millennium*, G. Lakemeyer and B. Nevel, eds. San Francisco: Morgan Kaufmann Publishers, 2002, pp. 239–269.
15. Pearl, J. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Francisco: Morgan Kaufmann Publishers, 1988.
16. Cowell, R. Introduction to inference in Bayesian networks. In *Learning in Graphical Models*, M. Jordan and J. A. Thomas, Eds. Cambridge, MA: MIT Press, 1998, pp. 9–26.
17. Shafer, G. R., and Shenoy, P. P. Probability propagation. *Annals of Mathematics and Artificial Intelligence,* Vol. 2, (1990), 327–352.
18. Lambert, D. A. A unification of sensor and higher-level fusion. In *Proceedings of the 9th International Conference on Information Fusion*, Florence, Italy, 2006.
19. Bowman, C. L., and Steinberg, A. N. Affordable non-traditional source data mining for context assessment to improve distributed fusion system robustness. *Proceedings of SPIE,* Vol. 8756, (2013), 875603.
20. Steinberg, A. N., Bowman, C. L., Blasch, E. Morefield, C., Morefield, M., and Haith, G. Adaptive context assessment and context management. In *Proceedings of the 17th International Conference on Information Fusion*, Salamanca, Spain, 2014.
21. Waltz, E. *Knowledge Management in the Intelligence Enterprise*. Norwood, MA: Artech House, 2003.

Alan N. Steinberg is recognized internationally as one of the leading experts in sensor data fusion and information exploitation, with over 35 years' experience as a designer, developer and operational user of major targeting, electronic combat, and intelligence systems.

- He currently enjoys the life of an Independent Consultant; having recently served as Principal Research Scientist at the Georgia Tech Research Institute. He held previous positions with Lockheed, Litton, TASC, USU/SDL, ERIM and CUBRC.
- A long-time member of the JDL Data Fusion Group, he has been active in refining the well-known JDL Data Fusion Model.
- In 2003 he received the Mignona Award for outstanding achievement in data fusion and has served on blue-ribbon panels for the US Government to evaluate and recommend technology developments and the restructuring of the Intelligence Enterprise.
- His current research involves developing techniques for adaptive context discovery and exploitation. Other recent work includes technology evaluation, forecasting and road mapping (2010-11), state-of-the art evaluation of Automatic Target Recognition techniques (2009), sensor and information fusion developments for Air and Missile Defense and various sensing and signal processing techniques.
- He has presented numerous courses, seminars and workshops in the US, UK, Sweden, Norway, Germany, Bulgaria, Armenia, Estonia, Australia, Singapore, China, South Africa and Turkey.