Toward an Understanding of the Cognitive Aspects of Data Fusion

Daniel M. Davenport
Raytheon Systems Company
dmd@hrb.com

12/14/98

1. Introduction

Data fusion is a key driver for achieving Information Superiority. Improvements in processor speed and visualization techniques, virtual reality and sensor technology are important drivers as well. But the bandwidth between humans and their machines is limited---the source being the form of human input/output we are restricted to and the natural speed of processing we perform at. Neither of these obeys Moore's law. Most of the significance of the sensor data that we collect will be lost if it is not interpreted for the human user before hand. Improvements in speed only make this gap wider. In an attempt to narrow the gap, new HCI technologies, such as virtual reality, are being developed. But this introduces a new problem, namely, how to display fused sensor data to the user in a meaningful way. So the gap remains. The present scenario looks a bit like Figure 1 where it is hoped that very intelligent software will automatically display exactly what we want and need and in a manner we can immediately grasp.

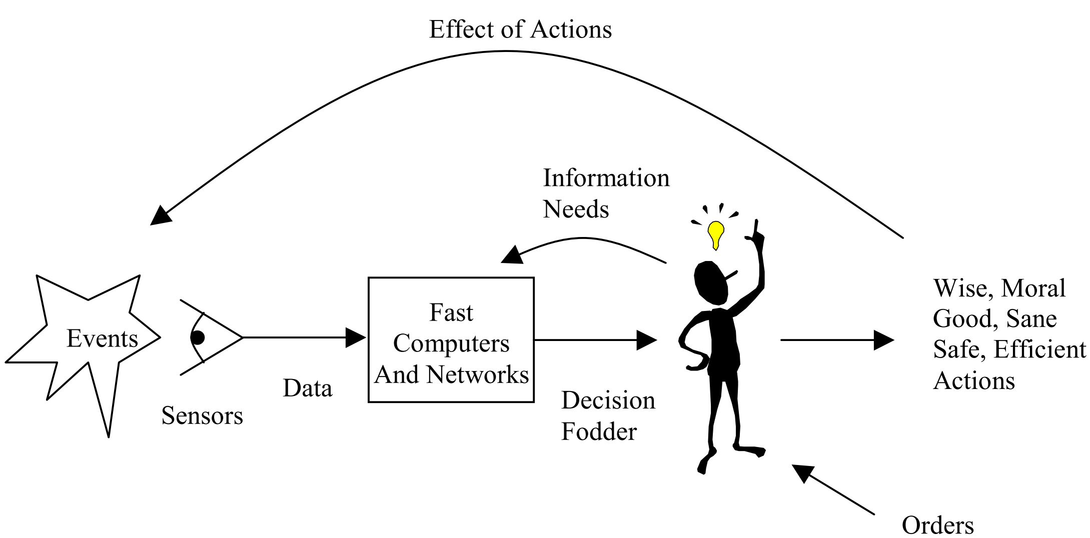

Figure 1. Making Wise Decisions

Form SF298 Citation Data

Report Date ("DD MON YYYY") 15121998	Report Type N/A	Dates Covered (from... to) ("DD MON YYYY")

Title and Subtitle Toward an Understanding of the Cognitive Aspects of Data Fusion	**Contract or Grant Number**
	Program Element Number
Authors Davenport, Daniel M.	**Project Number**
	Task Number
	Work Unit Number
Performing Organization Name(s) and Address(es) Raytheon Systems Company	**Performing Organization Number(s)**
Sponsoring/Monitoring Agency Name(s) and Address(es)	**Monitoring Agency Acronym**
	Monitoring Agency Report Number(s)
Distribution/Availability Statement Approved for public release, distribution unlimited	
Supplementary Notes	
Abstract	
Subject Terms	
Document Classification unclassified	**Classification of SF298** unclassified
Classification of Abstract unclassified	**Limitation of Abstract** unlimited
Number of Pages 12	

The idea, of course, of giving someone information is to allow him or her to make wise decisions. Note that, in this illustration, the user, after requesting the data, is essentially passive and uninvolved until *after* the information is displayed. That is, this way of thinking about the problem is essentially data driven instead of need driven. Granted, systems are designed for a particular purpose and it is assumed that when the system is put into operation that it gives the user what is needed in order to do the job. But, even when a user is granted the right to affect what data is presented, the methodologies driving the content of the data are still essentially fixed. This seems to be a waste of good computing power, the computing power of the human mind.

There is another point to consider here. Sensor data is of a fundamentally different type than non-sensor data. Sensor data comes to us in the form of random numbers in some distribution. This is what data fusion is best at, pulling information out of such distributions. However, it is increasingly necessary, indeed, ultimately necessary to fuse non-sensor data with sensor data. Non-sensor data, such as HUMINT reports, have an essentially different source than sensor data. Sensor data arises out of the natural processes of physics whereas non-sensor data arises out of the human mind, perhaps ultimately subject to the laws of physics, but clearly not in a simply distributed manner. Consequently, the state of the art in natural language processing is stone age compared to the state of the art of signal processing. At least in terms of results. The source of the former is simply of a different genre than that of the latter. If it is random at all, it is not random in the same way that physical processes are random. Hence, if any real progress is to be made in data fusion of non-sensor data, the source of that data needs to be better understood. This is essentially the conclusion of Hall and Llinas in [1] when they comment on the state of the art in Level 2 and Level 3 data fusion as follows:

> The main challenge in this area is the need to establish a viable knowledge base of rules, frames, scripts, or other methods to represent knowledge that support situation assessment or threat assessment. Unfortunately, there exist only very primitive cognitive models for how humans accomplish these functions.

The purpose of this paper is to present a cognitive model for how humans accomplish data fusion functions at a cognitive level along with some basic proposals for how to mimic these functions in software. To a limited extent we have validated our model by building a proof-of-concept natural language processing engine yielding a performance of 98% precision at 80% recall with reasonable processing speed ([2]). However, we will not burden the reader with technical details in this paper. Our purpose here is to provide a conceptual framework for implementation rather than implementation details.

We start with a model of human decision making since the making of wise decisions is, after all, the purpose of gathering and disseminating information. We present a model of computation in the next two sections that we call *Model Fusion* which derives from our reasoning about the human mind. We then state some implications all of this has on the discipline of data fusion. We close with a summary and conclusions.

2. Human Decision Making

Naively speaking, human beings engage in two distinguishable activities when coming to a decision. First, they assess the state of the world in which they are operating, and then they make choices among alternative actions to take, inactivity or delay of action being among these. We call the first activity *declarative* and the second *imperative*. That is, we come to some (declarative) conclusion, and then take some (imperative) action. Our actions are imperative in the sense that we typically insist on attempting to do only what we decide we should do in order to achieve some end state, goal or value. In terms of a computational model, declarative statements are often used to model how we *understand* the world and *rules* are often used to model how we respond to it.

According to modern cognitive psychology, the declarative side of our nature is often delineated according to what is *perceptual* and what is *conceptual*. The former being closely tied with levels one and two data fusion, the latter more closely correlative to levels three and four. At the same time there is a good deal of evidence from the physiology of the brain that the actual making of decisions is strongly influenced by our "emotional" side as manifested in our consciences and instincts. Our perceptual and instinctual sides are governed largely by unconscious processes. Our conceptual and ethical sides we seem to have more control over, that is, we can consciously affect our own thinking and make rational and wise decisions. We illustrate this in Figure 2. In at least an intuitive sense, these things taken together constitute judgement, the basis of wise decisions.

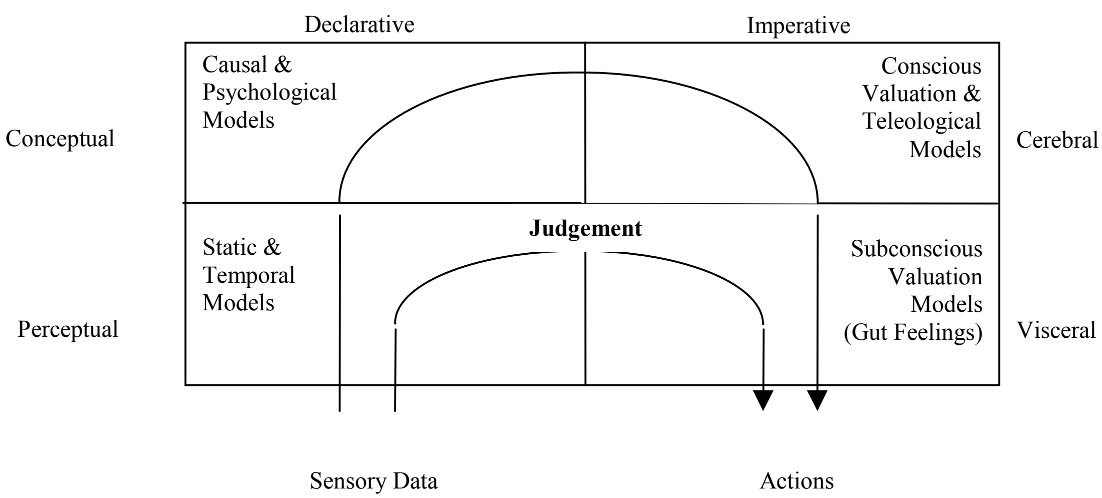

Figure 2. A simple model of decision making

The alert reader will note that there is little mention of the role of emotions in decision making except at the visceral level. This is because we prefer to think of emotion as arising out of the degree of disparity between what we *want* the world to be like and how we *experience* it. This disparity is given in terms of valuation models which measure the distance between what *should* be the case (as we value and desire it) and what *is* the case. Thus, we view emotion as an *effect* of our understanding and valuation of the world, rather than a *cause* as is sometimes tacitly held. Assessment of what the world is like and valuation of its significance, therefore, are basic to decision making in our model.

There is a great deal of evidence from the physiology and mapping of the human brain to support our model. Indeed, the same centers responsible for assessing the world and thinking about it are also centers of emotional responses. Also, the interconnection between the cerebrum and the portions of the brain controlling “lower level” mental processes are extremely rich. For example, it has been discovered that a lesion between the *prefrontal lobe* and the *amygdala* leaves the victim’s native intelligence intact, but renders them unable to make even the simplest decisions [3]. The *prefrontal lobe*, of course, is where our highest cognitive functions are centered. The *amgydala* works together with the *hippocampus* to regulate memory. But recall is not a problem with these people. Rather, the problem seems to lie with the manner in which memories of sensory data is stored in the first place. Damage to the *amygdala* inhibits the ability to recall information and damage to the *hippocampus* inhibits the ability to store new information. These structures work together to evaluate and store memories worth storing. The *amygdala* provides the value system (basically, how one feels about an experience) required to make this judgement. Thus, it appears that at least in part, the trouble is caused by the victim’s inability to correctly evaluate his experience. Of course, damage to this area due to the (now rejected) practice of frontal lobotomy, which cause lesions more in the center of the prefrontal lobe, results in affective flatness, that is, lack of a show of emotion. In our model, this is another indication of an inability to properly evaluate experience. Hence the absence of emotion.

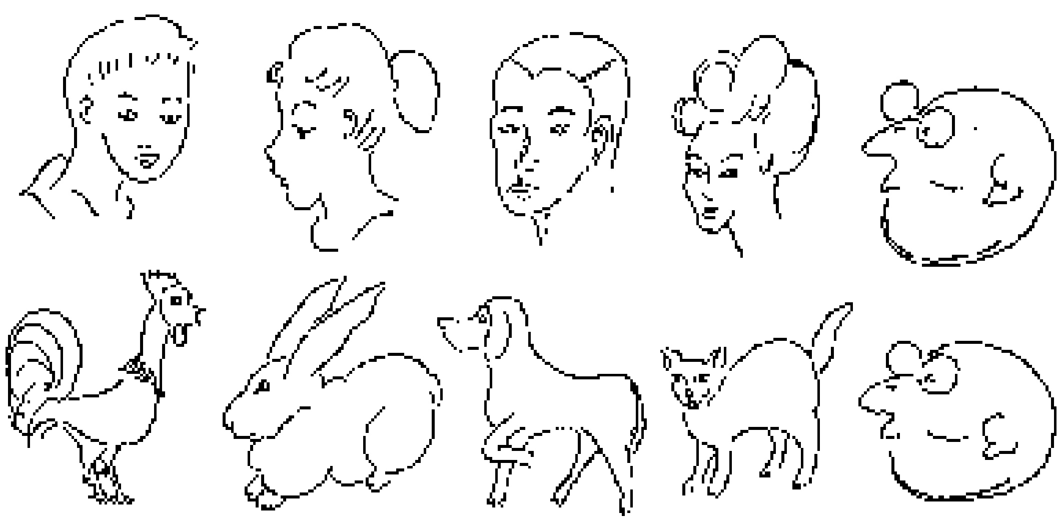

Figure 3. Is it a man or a rat?

We can illustrate the effects of values on perception as follows. Consider the pictures in Figure 3. If you cover up the bottom half and ask someone if they can see a man with glasses they will immediately point to the last figure on the right. However, when you cover the top figures instead, they will see it as a rat. In fact, many when shown the top and then the bottom are not able to see the rat until you point it out to them. This is particularly effective with someone who trusts you or is expecting you to give them accurate information. We claim that those that are unable to immediately see the rat experience a disconnection in valuation. According to our model, the subject is "set up" to see the figure as a man with glasses, in part by their expectations that you are giving them useful and accurate information. Their teleological model (i.e., their model of what giving information is properly *for*), along with the corroborating evidence of the pictures of other people, biases them to interpret those wavy lines as a man with glasses. That is, they subconsciously value the interpretation as a man with glasses over the interpretation of it as a rat. However, when it appears in a context that allows it to be interpreted as a rat, people often experience discomfort (confusion) or sometimes laugh. What they experience "out there" in the world is no longer in congruity with their current set of values regarding it. So emotions are inevitable. This phenomenon is known as *cueing* in educational psychology [4], where it is used to enhance learning by leading students to correct conclusions. It is also used in propaganda very effectively to mislead.

Thus, we can account for many of the factors that are known to be important to decision making. The effect of stress is to bias our system of values toward self- preservation. In fact, tunnel vision can result in extreme cases---a dramatic devaluing of peripherally visual phenomena. Method of presentation, including cueing and order of presentation [5], cause a great deal of variability in how we evaluate what data is important and how to interpret it. In the case of natural language this can be illustrated with the following example. Suppose I say "the dog bit John and then he ran away." Who ran away? Most people in our experience say the dog. On the other hand, if you say, "John was bitten by the dog and then he ran away," even though it is a bit awkward, most people (in our experience) say it was John who ran away. According to Ronald Langacker [6], this is due to the fact that we evaluate the first noun in a sentence as most important (*salient* in his manner of speaking) so that the pronoun "he" is interpreted as referring to the first noun rather than the more temporally proximate second noun. Thus, the values that we assign to our experience profoundly affect our understanding of the world.

So how does this affect how we compute things as human beings? First, we claim that, just as incorrect values and valuation procedures affect our judgement negatively, even so, correct values and valuation procedures are required for correct judgement. At least this is what our model suggests. Daniel Goleman, in his book *emotional Intelligence* [3], presents some compelling evidence to this effect. Hence, our values provide a kind of top-down pressure on our conceptual and perceptual processing systems that enable us to chose (through evaluation) what data or experiences are important. The functions of the *amygdala* and *hippocampus* relative to memory also suggest this. Storage and retrieval of memories (and hence what we "choose" to incorporate into our experience) are in part a function of how we feel about them and, consequently in our model, how we evaluate them.

Second, we claim that values are intimately connected with very general and vague memory structures that we call *quanta of meaning* or *quanta* for short. There is ample evidence that the *amygdala*, for example, helps us recall specific events in our lives with high emotional content, especially frightening ones. Also, the *temporal lobes* allow the access of declarative memory, that is, the conscious recall of what was learned as well as when and where it was learned. Significantly, it also functions to attribute personal meaning and significance to the constant stream of sensory input we experience. We argue that it is reasonable to model this latter phenomenon with quanta, which represent schematic or "average" experiences in the form of general rules of thumb for evaluating the significance of one's current situation and experience.

3. Modeling the World with Quanta

In this section, we will apply what we have asserted about the human mind to creating models of the world with quanta. First, however, we want to give the reader more intuition as to what we mean by "quanta." To do this we start with a discussion of the properties of the information we gather so we can be sure to capture these properties with our knowledge structures. First of all, information about the world is often ambiguous and inconsistent. Hence the use of statistics and fuzzy logic model this. It is context dependent since what we want the information *for* and how we value it are just as important to how we interpret it as is its "objective" content. Indeed, information that we do not value at all is ignored. Also, at best, we only ever possess an abstraction of what is "out there," stripped of its full context within the totality of the world and history. We therefore not only interpret information in the context of what we need it for, but we also strip it of its full context in the process.

In order to restore a part of that context, which we cannot fully comprehend, we relate to the world through quanta and mental pictures that "fill in the blanks," as it were. We don't simply construct the "truth" out of atomic brute facts using logic. It *may* be the case that at a neural level we do. However, in practice, we use the summation of many experiences to interpret the world correctly. This summation, we claim, is in the form of a vaguely defined set of mental states that constitute a sort of data structure that captures simultaneously particular experiences as well as summations (generalizations) of them in hierarchies. These hierarchies, we claim, are indexed not only by their sensorial properties, but also by how we value them in themselves and relative to other related experiences as they bear on what kinds of goals we have.

Concerning this last point, we want to be very clear that we believe that goals are constructed from the fusion of many quanta encoding our experiences. That is, we choose our goals according to what we *value* and not the other way around. Truly intelligent systems are not only able to realize *goals* (as Newell [7] claims is sufficient), but to realize *values* which change with circumstances. Quanta and the values that they encapsulate combine to form a flexible picture of both a set of means and a set of end states that are desirable to obtain. Thus a goal is born.

In order to fuse quanta they must have an internal structure that allows their corresponding substructures to be fused recursively. To describe this we borrow heavily from cognitive psychology and the work of Ronald Langacker in Cognitive Linguistics [6]. The notions of figure and ground are familiar to anyone who has seen a picture like the one in Figure 4. When the inside is seen as the ground then the figure of two faces emerges. When the black outer part is viewed as the ground, then the figure of a vase appears.

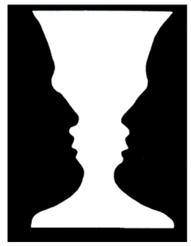

Figure 4. Figure and Ground

The *ground* is, roughly speaking, the context in which the figure is interpreted. The figure is the part of the ground that is brought to prominence, or in Langacker's parlance, to *salience*. The figures within the ground give a partial inventory of the kinds of things and relations that exist in the ground. To give a linguistic example, the assumed context that you give to the sentence, "a cat ate the rat" is the ground of the meaning of the sentence as you interpret it. The figures are the cat, the rat, and the relationship they bear to one another in the eating of the rat by the cat. In the face/vase example, the relationship that the two faces bear to one another, namely, that they are facing each other, is a part of the figure, not the ground. When we pick the figures out of a ground (that is, bring them to salience) we define the whole scene intensionally (that is, according to its properties, in this case, its components and their relationships within the ground). We claim, generalizing Langacker's claim, that all human knowledge structures (that which we have dubbed quanta) have similar components. They each have a ground, representing all the potential or tacit elements within it that can be brought to salience, as well as elements that are currently salient and therefore within our current conscious experience. The tacit elements of the ground constitute a very ambiguous definition of it. The figure gives it more a precise definition. The degree to which an element within the ground is salient determines the degree to which we value it in the current situation. When we are highly focused, the figure is relatively sharp. When we are more dreamy (what philosophers call being in a pre-theoretic mode of thinking), the ground returns to an ambiguous state. This internal structure, along with the values within it, constitutes the *intension* of a quantum. The ground, its figure, and its value relative other grounds possessing salient figures, all constitute a quantum knowledge structure.

Quanta can be thought of as points whose components are the tacit elements of the ground and its figures those elements that have a positive weight. The position of an element within these points, which amounts to naming, provides a method for accessing each component. Another way of thinking about quanta is as a space in which many objects exist in various relationships (some temporal perhaps) but only a few of which are currently considered to be important. In either case, the important thing to note about quanta is that the elements within their grounds can be quanta themselves. In this sense, quanta generalize classes in object oriented languages.

4. Computing with Quanta

Two quanta with mutually consistent internal structures (that is, similar in intension) are comparable by comparing their corresponding components, just like points. For example, the sentence "something did something to something" represents a very general transitive event. This has a comparable structure to "the cat ate the rat" which is an instance of the former. Notice, however, that the latter sentence *is still ambiguous* since we know nothing about the cat, the eating, or the rat. Thus, "Boots gobbled up my pet rat" is an instance of both the former. Rather than saying that the last sentence "is-a" instance of the others, we prefer to use the language of definitions and say that the latter *extends* the former two. Thinking of quanta as spaces is instructive here. The first sentence represents a relatively large subspace of a possibility space. The latter two are progressively smaller subspaces in the same space, nested one with the other and both within the first. Their point-like qualities allow us to make meaningful correspondences between the figures in these grounds. This, in turn, allows us to index them within an extensional hierarchy. The astute reader will also notice that there are two sub-hierarchies from which this larger one is constructed, namely a hierarchy of nouns and a hierarchy of verbs.

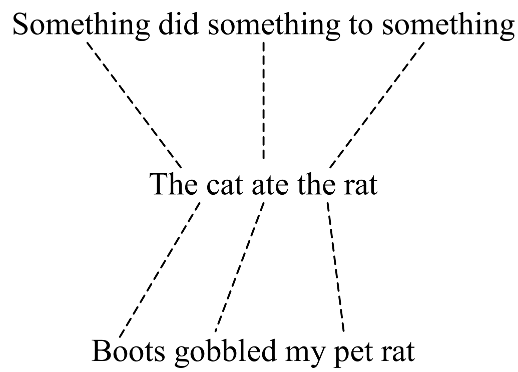

Figure 5. Extension of Concepts

Based on the notions of intension and extension, we can describe a database that uses the index that intensional elements of a ground provide to represent directly the extensional hierarchy implied by it. Note that if we start at the top of the hierarchy with the most vague quanta at the top, a kind of summation of all those below it, that we can identify whether or not a quantum lives within the hierarchy by comparing it to this top element. If it is an instance then we know it has a place somewhere in the possibility space defined by the top element. This is a natural way of ensuring database consistency. If we proceed down the hierarchy to the next level of quanta, then we can repeat the process as long as exactly one quanta at that level is extended by the quanta we are categorizing. The quanta we are categorizing will drift down to its natural level in time linear to the depth of the hierarchy, typically logarithmic in the number of entries. That is, we can categorize quanta quickly.

By the same token, we can use quanta as queries by categorizing them and retrieving all the entries below them (that is, that are instances of them). This process is linear in the depth of the hierarchy and the number of database hits. Thus we can retrieve stored quanta efficiently as well. This is illustrated in Figure 6.

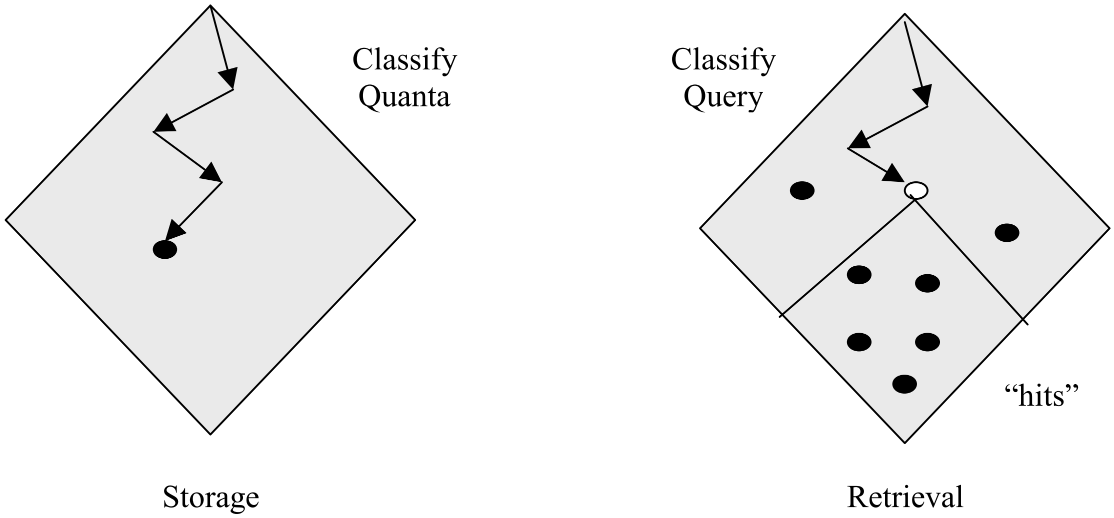

Figure 6. Storing and retrieving quanta

We can combine quanta that are not completely comparable at the points at which they do agree. This process we call *fusion*, though it is fusion conceived of at a more abstract level. For example, consider the uncertainty spaces of a pulsed radar and a infrared imaging sensor. When these two spaces are correlated correctly, their intersection is a much smaller uncertainty space (see [1]). These spaces require particular methods to fuse them together. We can also combine linguistically the concept of *cat* (which has a tacit eating component) with that of *rat* and of *eating* and their respective tacit components. The correct correspondence is made through the rules of English grammar. The corresponding (tacit) elements of each are individually fused, fusing the scene into a whole. According to the rules of cognitive linguistics, the most salient figure in this ground is the cat, then the eating, then the rat. Other concepts, like *where* the cat dined, remain tacit.

Thus we can fuse quanta together by fusing corresponding elements in appropriate ways along with their tacit elements which become salient as new information is made available. Note that this implies that the meaning we assign to symbols or data that are presented to us in different orders is often different since the order determines their salience and salience is a part of how we interpret symbols. This is consistent with what is already known about the effect of order of presentation on decision making [5]. Hence, we can construct less and less ambiguous scenes of the world out of very ambiguous quanta but at the cost of becoming biased by the order in which they are presented. We hypothesize that parallel processing in the human brain ameliorates this to some degree.

One might now ask how we know what quanta to fuse in the first place and how to access them. First, recall that we can order quanta in extensional hierarchies. Thus, in principle at least, and often in practice, we can index quanta according to their intensional properties as well as the salience values of their components. This is consistent with the manner in which the human brain is thought to index memories, as we have seen. When a symbol is presented to us that we are familiar with, it brings to mind a general quanta which is combined according to some convention or grammar with the quanta of other symbols around it in some ground space (temporal or spatial or both). The net result is a quantum representing all the possible meanings of the symbols we have viewed with a preferred meaning encoded in its figure, that is, in the elements of the ground that have become salient during the fusion process. In the case of linguistics, words in a particular temporal order are mapped individually to various quantum structures and are combined sequentially to form a coherent picture of the meaning of the sentence. For example, the subject of each English sentence is always the most salient and therefore the most influential in determining the meaning of subsequently spoken words. That is, the cognitive ground for interpreting the next sentence is composed of the current state that the mind is in. In terms of our model, that current state is a figure against a ground as represented by a quantum data structure. Subsequent information is interpreted in terms of this structure.

5. Implications for Data Fusion

The concept of quanta, as we have presented it, generalizes what is meant by a possibility or uncertainty space. Thus, they provide a principled basis for correlating and fusing sensor data with non-sensor data so that they can be fused into a coherent interpretation of the world. In addition to this, they provide a of narrowing down, through the top down methods we have introduced, the data that needs to be looked at in order to accomplish data fusion. These top down methods are drawn from a study of how the human mind accomplishes the same thing. Hence they are subject to many of the same problems related to human interpretation of objective facts. However, these methods also compliment the inherently bottom up methods of sensor data fusion, providing a natural means of supplying needed interpretive context to the lower levels of data fusion. At the same time, however, without these tried and true bottom up methods, all the top down methods in the world will avail nothing, lacking the objectivity that bottom up methods supply. We conclude, therefore, that both our top down methods as well as bottom up methods are required to complete the success of data fusion in enabling practical information superior systems. We believe our notions and methods will be useful to creating systems robustly capable of performing levels two and three data fusion.

We want to note here that our methods are not incompatible with current symbolic methods, such as case based reasoning and black board technologies. At the same time it is not incompatible with existing biological methods, such as neural nets. In fact, we strongly suspect that a judicious combination of these methods, integrated along the lines we have proposed, will ultimately be required to build truly robust data fusion systems.

At the same time, we also believe that our methods, when properly combined with existing methods, can assist in achieving robust level four fusion as well. For one thing, goals can be constructed out of quantum structures representing desirable (i.e. highly valued) end states and means. Specification of the means to the ends allows the system to determine the specific source requirements of lower levels. Under the assumption that lower level functions are able to communicate to level four functions their capabilities and limitations, level four can then identify what is required to improve system performance. For example, if the system knows it needs a certain resolution to perform a certain task but none is available, it can ask its human masters to build the modules it needs.

Our model also has implications for human computer interfaces. Busy people prefer to have information displayed to them in familiar forms. For example, a commander requires information to be displayed in operational terms, not technical terms endemic to how the information was gathered or how it was processed. Quanta give a basis for decomposing user requirements, as expressed in terms of their goals or missions, into the language of gathering and processing. The user's information requirements, preferences, and rationale for wanting the information can be represented and fused together into a coherent quantum structure representing the possibility space that is of interest. The most salient portions of the quanta are given higher priority. The request for information is broken down by the salient components of the quanta (which are themselves quantum) into other information requests at a lower level. These in turn are similarly decomposed into their components. At some point highly specific requests will be made to well defined processes that identify themselves as capable of supplying the needed information. They supply the information along with a rationale for why they got it. This rationale is fused along with the information retrieved back up through the calling quanta providing the user with an answer. Only the most salient items need be displayed, not all the details. If more information is required the user can drill down into the natural components of the top layer quanta and get a rationale for the conclusions that the system drew along with a breakdown of the information into its most salient components. Providing a rationale is necessary to true drill down.

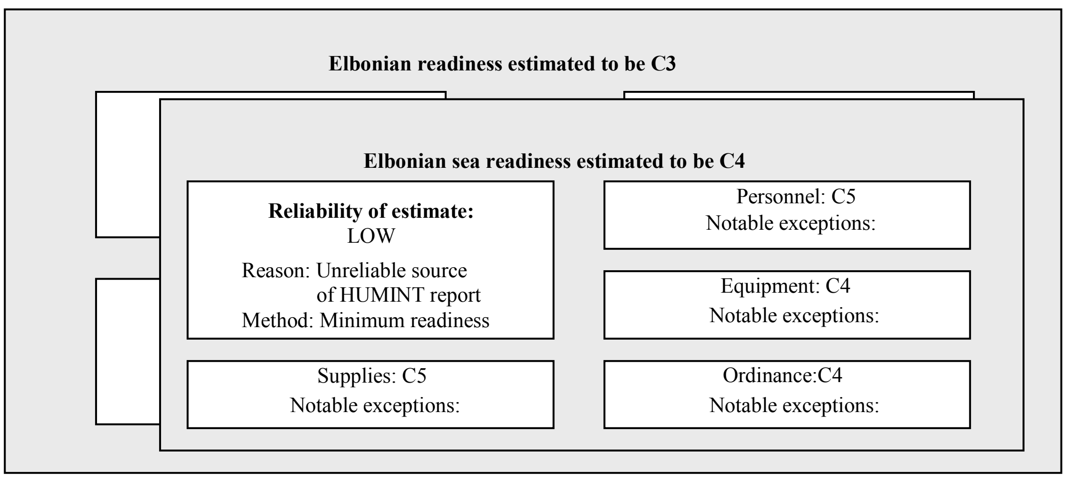

Figure 7. True drill down provides a rationale.

6. Summary and Conclusions

In this paper we have presented a simple model of human decision making that predicts that processing information to make decisions requires valuation of the data of which it is comprised at all levels. We have seen that both top down and bottom up processes are used in human information processing. We have noted that information about the world is ambiguous and often inconsistent and therefore requires data structures that are able to represent this uniformly whether from sensor or non-sensor sources. We introduced the notion of quanta which captures the notion of figure and ground, a necessity in representing ambiguous information. We also noted that quanta have both a point-like structure, for referring to components, and a space-like structure, along with a measure of the relative importance of the components contained within the ground. We noted that we could index these structures by their components, that is, by their intension. This yields a natural ordering of these structures into an extensional hierarchy which decomposes into smaller extensional hierarchies. We noted that this gives rise to natural methods of query and retrieval of information encoded in quantum form. We also noted that we can fuse quanta representing symbols to form an interpretation of the symbols. Thus we can, in principle, fuse sensor data with non-sensor data in a uniform manner. Finally, we concluded that our model has the potential to be useful to levels two, three and four data fusion as well as to the implementation of human computer interfaces.

7. Bibliography

[1] Hall, D. L., and Llinas, J., An Introduction to Multisensor Data Fusion, proceedings of the IEEE, Vol. 85, No. 1, January 1997.

[2] Davenport, D. M., and Jennings, M. A., The ConceptCrawler Web Search Engine, CIIP Technical Report (available from the author), May, 1997.

[3] Goleman, Daniel P., *Emotional Intelligence,* Bantam Books, 1995.

[4] Dwyer, F. M., *Strategies for Improving Visual Learning*, Learning Services, Box 784, State College, PA, 1978.

[5] Llinas, J., Drury, C., Bialas, W., Chen, A*., Studies and Analyses of Vulnerabilities in Aided Adversarial Decision Making*, SUNY at Buffalo, Dept. of Industrial Engineering, Feb., 1997.

[6] Langacker, Ronald W., *Foundations of Cognitive Grammar*, two vols., Stanford, Calif. Stanford University Press, 1987-1991.

[7] Newell, A., the Knowledge Level, *Artificial Intelligence Magazine*, 2(2):1-20, Summer, 1981.