A Framework for Network Visualisation: Progress Report [1]

M. M. Taylor
Martin Taylor Consulting, 369 Castlefield Avenue, Toronto, Ontario, Canada M5N 1L4
mmt@mmtaylor.net

On behalf of the IST059/RTG-025 Working Group on *Framework for Network Visualisation*
J.-T. Bjørke (Norway), M. R. Nixon (USA), M. M. Taylor (Canada), A. K. C. S. Vanderbilt (USA), M. J. Varga (UK)

FRAMEWORK FOR NETWORK VISUALISATION

Hundreds of different representations of networks have been produced for different purposes. Many are very well suited to their purpose, but it is seldom clear how to generalize their insights to other situations. Networks can be described in many different ways, each representing an attribute of the network that might be useful if displayed for some user task, of which there are an indefinite number. The objective of a Framework is to provide some structure in this universe of possibility, in order to assist a user with a particular task to find a good way of displaying useful network attributes, or to help a designer produce a display that will help a range of users with a variety of related tasks. The IST-059/RTG-025 Working group on Framework for Network Visualisation has developed a conceptual structure for a Framework, based around two previously developed reference models for visualisation, a set of possible local and global dimensions of description of networks, the concept of "embedding fields" both for networks and for displays, a typology or taxonomy of data and display types, and a categorization of user task types. The resulting structure will be developed, along with the complementary IST-059/RTG-025 Survey of applications for network visualisation, into a structured guide for users, researchers, and developers.

1.0 FRAMEWORK BACKGROUND

Users of computer-based displays for network visualisation have no particular need to see "pretty pictures", but have some problem in the real world with which a computer-aided analysis may help. Often, the real world issue concerns relationships of one kind or another. When there are more than two related entities, the issue involves a network. That network should be displayed to the user in some manner or other, but different users might want to see different things about the same network, and different kinds of networks offer different things to be displayed. The result is a proliferation of applications that give wonderful pictures, but there is very little help available to aid a user to discover what kind of application and display will be useful for the real task. Hence the need for a Framework for network visualisation.

1.1 The Nature of a Framework for Network Visualisation.

Several stages of abstraction intervene when computer support helps the user address a real world issue that involves a network, (Fig. 1). First, some of the data relevant to the networked entities is obtained and stored in the computer. Second, mathematical algorithms operate on those data to abstract a set of properties of the network. These properties constitute the universe of what might possibly be displayed to the user. Next,

[1]

Taylor, M.M. (2006) A Framework for Network Visualisation: Progress Report. In *Visualising Network Information* (pp. 3-1 – 3-22). Meeting Proceedings RTO-MP-IST-063, Paper 3. Neuilly-sur-Seine, France: RTO. Available from: http://www.rto.nato.int/abstracts.asp.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
01 DEC 2006	**N/A**	**-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
A Framework for Network Visualisation: Progress Report	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Martin Taylor Consulting, 369 Castlefield Avenue, Toronto, Ontario, Canada M5N 1L4**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM002067., The original document contains color images.

14. ABSTRACT

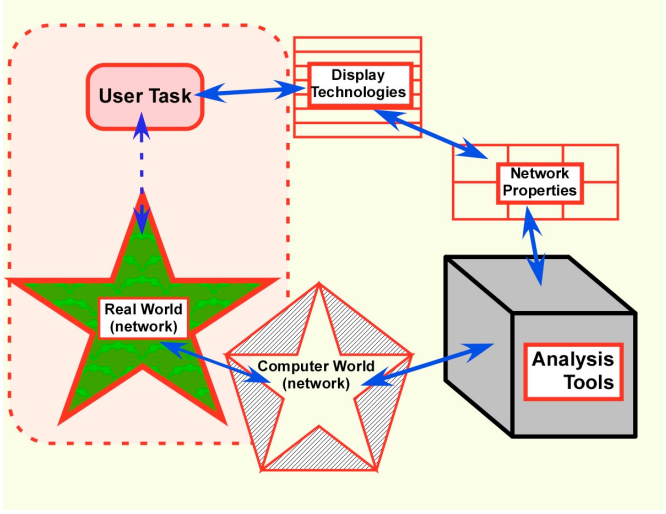
15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**UU**	**60**	

display algorithms and apparatus operate on a selection of the displayable properties to create a display presentation. Finally, using the presentation, the user may visualise the real world situation of interest.

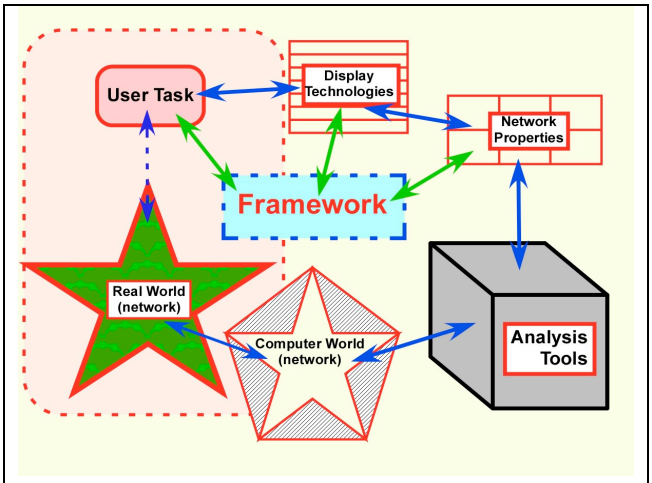
The transformation and abstraction processes all have many different possibilities, some of which may be under the user's control, and some not. The abstraction of data from the

real world into the computer probably is not under the control of the user, so the Framework probably should ignore that process. The black box of analysis tools has a somewhat independent existence, in that if network data are available, any appropriate tool may be chosen, and its output is a property of the network. Because of their independence, no added benefit is likely to come from simply incorporating the list of available tools into a Framework. But after that, the abstractions and the processes interact with each other, and it is here that a Framework may prove valuable. Different network properties are likely to be well suited to different display technologies, and the user's ability to visualise is very much affected by the manner in which different properties are displayed.



1. Abstractions in network visualisation

Over the years, a great number of different display techniques have been developed to show off different kinds of network properties. Many of these work very well, and many are attractive to see, though the two classes do not necessarily overlap very closely. Some work well for networks with only a few nodes (vertices in the mathematical abstractions) and links (mathematically, edges or ties) but degenerate into an unintelligible mess for larger networks. They do not scale well. Some may express different types of nodes in ways the user can visualise, whereas for others, a node is a node is a node. Whatever their strengths and weaknesses, almost all have been developed *ad hoc*, for one particular application.



2. Where the Framework fits.

One of the objectives of a Framework is to provide some kind of language or structure that would relate display techniques to the kinds of network properties for which they are suited. The other side of the problem is equally important: to provide some kind of a language or structure that would relate display techniques to the kinds of visualisation they might facilitate. The Framework therefore can be seen as a structure, or as an intermediate language that would allow network properties, display techniques, and the user's visualisation abilities to be considered together (Figure 2).

A good Framework may also be used as a process or cookbook that would help the user with a network-related problem to choose the right tools to help in visualising a solution. Categorizing the implications of the problem for the choice of perceptual modes may assist the user to select the appropriate kind of display, and therefore the appropriate technology and the kind of network properties that could usefully be displayed.

The Working Group therefore conceives a Framework for Network Visualisation as having several possible roles:

- An interface that connects a network-related task requirement with the available display technologies
- An interface that connects the available display technologies with computed network properties
- A way of categorizing and describing user needs, display technologies, and network properties
- A help to users in assessing the nature of their requirements
- A guide to users in choosing a visualisation system suitable for their application need.
- A guide to developers and researchers as to unmet needs.
- A structure on which to hang the aspects of networks that might be displayed
- A structure on which to hang the likely ways network attributes might be used in real-world tasks
- A process or cookbook for linking tasks with network attributes
- A process for linking task–network attribute pairs with appropriate display technologies

1.2 Why try to create a Framework for Network Visualisation?

Why is IST-059/RTG-025 concerned with the task of developing a framework in the first place?

- If I have only a hammer, every job seems to require nails.
- If I need to fasten something, how do I know hammers exist?
- If I need something fastened and I know the tools exist, do I glue, screw, staple, or nail?

"I" would want a Framework that categorized fastening jobs in terms of what tools were best for those jobs, and categorized tools in terms of what kinds of fastening jobs they did best. "I" would like to be able to consider the task at hand in terms compatible with the capabilities of the tools, and would like to know how to get the kind of tool that seemed best suited to my job.

When it comes to providing displays that help users to visualise networks, there are several specific reasons for wanting a Framework:

- It is usually not clear how the insights that led to particularly effective representations can be generalized to new situations.
- A good Framework might help identify the conditions for which different insights are helpful.
- Users need to see different aspects of network structure and function, and some of those aspects are not well served by extant display techniques (they may have a hammer and a chisel, but have never heard of a drill).
- Users usually choose to see those aspects for which effective display techniques are available (they have a hammer and like to nail things together).
- A good Framework may help inspire research on new modes of display for different kinds of network properties (the need for drills may be made evident).

The user should be able to describe what it was about the network that she would like to see, to do so in a way that allowed her to find an application that would be able to extract that information, and that would allow her to display it in a way she could understand. IST-059/RTG-025 has a separate Working Group charged with surveying software for network visualisation. If the Survey and the Framework are effectively designed, a user should be able to use the Framework to analyze the task requirements, and using the results, to inspect the Survey to find a suitable application. The existing Survey may, to some extent, guide the detailed categorization of applications for the development of the Framework. Conversely, the more theoretically

based elements of the Framework may assist in structuring an effective Survey.

To summarize, A Framework for network visualisation should include:

- A structured approach to describing user needs
- A structured set of displayable properties of networks
- A structured way of describing display techniques
- A process to help the user match needs to displayable properties and appropriate display techniques.

No such Framework has yet been developed, though its outlines are clear. This paper details the progress of the working group towards that end.

2. FRAMEWORK FOUNDATIONS

The concepts being used by the IST-059/RTG-025 Framework Working Group stem from a variety of independent roots. These include, in no particular order:

- Reference Models for Visualisation
- Categorizng the modes of perception as applied to user tasks
- Dimensions of description of network types
- Mathematical abstraction of network properties
- Real world contexts of networks and the concept of "embedding fields".
- Taxonomies of data and display types

Several of these conceptual starting points come from the work of the predecessor groups of IST-059/RTG-025.

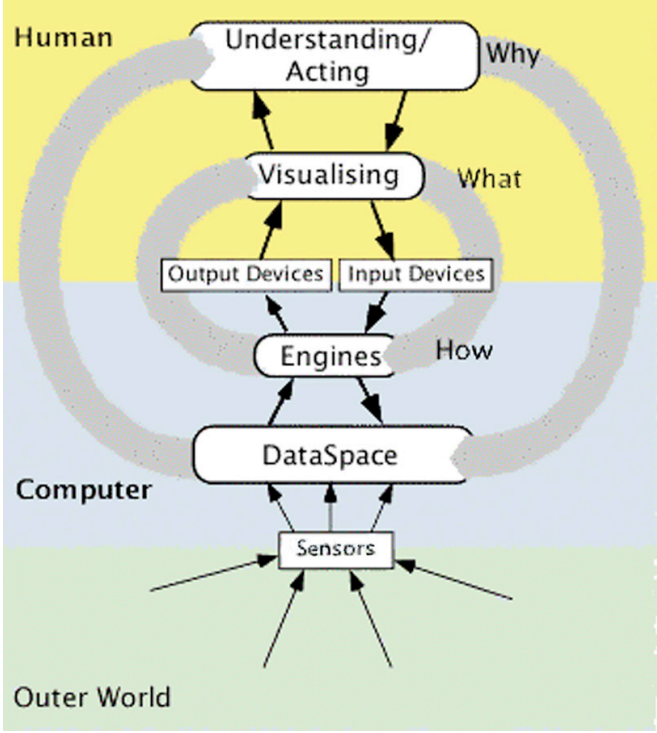
2.1 Reference Models for Visualisation

One conceptual basis for the Framework is provided by two independently produced but complementary reference models for visualisation. One, the "VisTG Reference Model", was developed by the predecessor groups of IST-059/RTG-025, whereas the other, the "RM-Vis Reference Model", was developed by a similar international defence group, the Technical Collaboration Committee (TTCP) C3I AG-3 Working Group, commonly known as "AG-Vis". The VisTG Reference Model is a functional model that considers the processes that allow the user to visualise data held in a computer to be a set of nested feedback loops, whereas the RM-Vis Reference Model is a descriptive model that offers a three-dimensional representation of visualisation applications. The RM-Vis three dimensions are domain context (who has the problem), what kinds of thing are to be displayed, and the visualisation approach (or display technology).

The two reference models have not yet been properly coordinated, but they seem to be mutually supportive, and each seems to have a place in defining an effective Framework for Network Visualisation.

2.1.1 VisTG Reference Model

The VisTG Reference Model is founded on the idea that analysis and design should always start with asking what the user wants. The whole reason the user resorts to computer support is that it provides a view on the world that is not easily obtained by direct observation. The user's objective, then, is usually to understand the implications of the contents of the data stored in the computer, and to act on those data or on the real world represented by them so as to bring them to a desired state. This objective is indicated in the grey outermost feedback loop in Figure 3.



3. The VisTG Reference Model

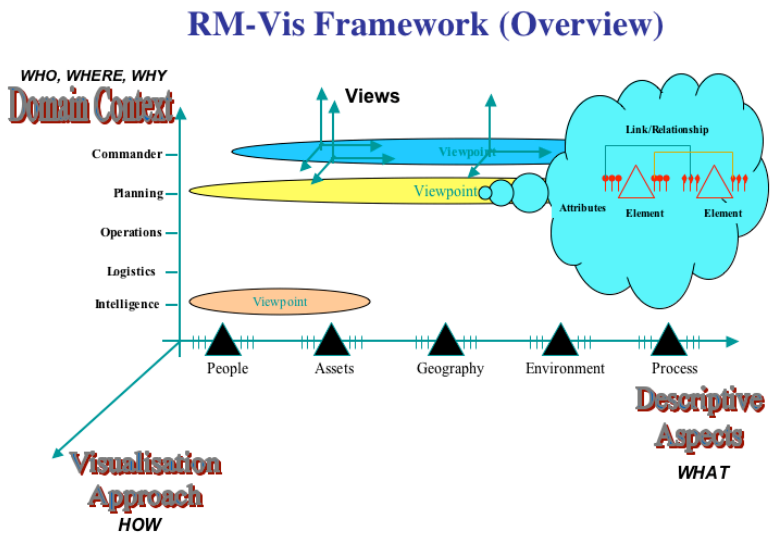
Visualisation is taken to be one of the routes to understanding, the other being logical analysis (which is not specifically represented in the VisTG Reference Model). To attain understanding is the user's reason for visualising; computer-based processes, generically called "Engines", are the means. The VisTG Reference Model therefore connects the visualisation processes in the human with the engines in the computer by way of an intermediate-level feedback loop. This loop implies the ability of the human to choose and to control the particular engines to be used. In the picture of Figure 1 or 2, the box marked "Analysis Tools" contains many of the engines; the block labelled "Display Technologies" contains others.

Just as the human cannot telepathically understand the implications of the contents of the dataspace, but must resort to visualisation and logical analysis, so the human's visualisation processes cannot telepathically communicate with the engines and must use physical sensors and muscles to communicate through input-output devices. The VisTG Reference model shows the actual connections by the solid arrows, in contrast to the grey loops that designate the conceptual feedback loops. When analysing or constructing systems to aid visualisation, the grey loops are the important ones that determine what must be done; the black arrows are the mechanisms whereby it is done.

The VisTG Reference Model is more fully explained in the Final Report of IST-013/RTG-002, a predecessor to IST-059/RTG-25 [1].

2.1.2 RM-Vis Reference Model

The RM-Vis Reference Model (Figure 4) is very different in character. It does not attempt to address the process of visualisation, but instead looks at the external circumstances. It provides a framework within which a particular visualisation task can be described. The RM-Vis model is usually depicted as representing a visualisation task in a three-dimensional space, one axis being the role of the user (e.g. Commander, Planning, Operations, Logistics, Intelligence), one being the display content (e.g. People, Assets, Geography, Environment, Process …), and the third being the approach taken to depict the thing to be visualised.



4. RMVis Model structure

Whereas the RM-Vis model indicates but does not make explicit the "Visualisation approach" dimension, the VisTG Model concentrates on that dimensions whil ignoring the other two (the user's role and the display content). There thus seems a good opportunity to enhance the framework to combine the two reference models into something that encompasses the useful features of both. This has not yet been attempted.

The RM-Vis model is more fully explained in the Final Report of TTCP C3I AG-3 [2].

2.2 Modes of Perception

A second conceptual basis for the Framework comes from the study of human perception. Perception is the human's window on the world. All we know about the world is in what we perceive. What we intend to do with that knowledge makes a real difference to how we gather it, and hence to the appropriate displays, when the perceived world exists in the computer's dataspace. Four distinct modes have been recognized [1, 3, 4], and were somewhat extended in [4].

The four basic modes of perception are:

- **Controlling/Monitoring**: The data are changing continuously, and the perception is tracking the ongoing changes. In "Controlling" mode, the person is acting to influence the thing being perceived, whereas in "Monitoring" mode the person is passively maintaining situation awareness without acting to influence the thing being observed. Monitoring can switch directly into Controlling, if the necessary effector mechanisms are available. Perceptually, the two are equivalent.
- **Searching**: For immediate purposes, usually in support of ongoing controlling or monitoring, some missing piece of data must be sought. Often this implies changing the field of view. If we are talking about observation in the real world, as opposed to a computerized dataspace, this amounts to sensor

redeployment. When found, the sought data are incorporated into the current situation awareness, possibly changing the perceiver's appreciation of the existing state of the world.

- **Exploring**: Exploring provides background information, by analogy to the production of a map. Data acquired by exploration are assumed to remain valid at a later time when they might provide context for data being monitored or controlled. Exploring reveals the structure of things such as networks. It is always performed at a lower priority than is ongoing controlling/monitoring or search, since those activities concern ongoing events. Exploring and Searching are often confused, as they may sometimes provide similar results. The difference between them is the difference between, on the one hand needing to write something and looking for a pencil you find in a drawer, and on the other hand wondering what is in the drawer, finding a pencil there, and remembering that the pencil will be in the drawer if you later want to write something. Memory is key to the difference between Exploring and Searching. Searching requires no memory for the thing found, since it is used immediately. The result of Exploration is assumed to be in its remembered state when it later becomes useful for some purpose.
- **Alerting**: Alerting has a status somewhat different from the other three, all of which admit the possibility of actively investigating the content of the observable world, but are restricted in the number and detail of independent entities being perceived at any one moment. Alerting systems are autonomous and are, in principle, unlimited in number and in their range of examination of the world. An individual alerting system consists of an algorithm that takes in data from known sources, and issues a signal indicating whether (and possibly to what degree) the pattern defined by the algorithm exists in the source data. It is a filter, perhaps a highly active filter, but a filter nevertheless. Any number of alerting algorithms may be applied to the totality of data that could be observed from all available sources. The function of the output of the algorithms is to provide the owner of the alerting system (human or computer) with an indication that it might be useful to divert attention to the area of data to which the algorithm was applied. Humans have several built-in alerting systems, such as the one that senses unexpected movement in the visual periphery. That system frequently results in a quick flick of the eye to the location of a sensed movement, to determine whether the movement signifies something that ought to be monitored, or whether the person should take some action. Alerting is autonomous and (even in humans) not a conscious activity.

In the context of network visualisation, any or all of the four modes may be important, but in many applications Exploring is the most important.

Network discovery is Exploration, and many tasks involving networks concern mapping the network or at least parts of it. This holds whether the network is of contacts among possible terrorists, the influences linking enzymes and proteins, the risks and remedies of infrastructure damage, or the political relationships in an area subject to peacekeeping operations.

Looking for the node with highest centrality in a network is Searching, a task that might be concerned with either an attack on, or the defence of, a network. Searching looks for something predefined; it is never serendipitous, though in the course of a Search, serendipitous discovery is possible

Tracking the network traffic among suspected conspirators is Monitoring, and doing that while inserting crafted disruptive messages into the traffic would be Controlling.

Filters that are set to respond to particular patterns of packet traffic and send messages to network monitors when they happen would be Alerting systems. Their actions are silent, and never displayed

to the user unless the filter algorithm produces a sufficient output.

The different perceptual modes have different implications for effective display, and those implications must affect the form of a useful framework. For example, Controlling or Monitoring usually benefits from a display that permits the user a real-time view of ongoing changes within a network terrain that has previously been specified. The pattern of changes should ordinarily be emphasised, rather than the stable structure of the net, which is necessary only to give some context. Exploring, on the other hand, is unconcerned with rapid changes, since those imply data that will not be valid when the time comes for the explored network terrain to be used. If shown at all, changeable aspects of the network probably should be shown only in a mode that indicates their unreliability for future reference. The display needs to show contexts that might indicate unknown regions that might profitably be examined. Exploration without the context of existing knowledge is seldom easy to use. Exploration is particularly likely to benefit from display of the embedding field(s) of the known parts of a network (see below for a discussion of embedding fields of networks and of displays).

2.3 Dimensions of description of network types

The basic classical network consists of a set of nodes (vertices), any pair of which either is or is not connected by a directed link (edge, tie). Some kind of traffic may pass from one node to another over a link. One might call such a network a "binary" network, as any node or link has only one attribute, its existence. Real networks are much more complicated, and the Framework must take that complexity into account. At this writing, the Framework group recognizes several dimensions of description, broadly divided into dimensions of description of the local properties of single nodes and single links, and dimensions of description of networks or subnets consisting of at least two nodes connected by a link.

2.3.1 Local properties of nodes and links

This classical concept of a network has been extended in several directions, most of which have implications for display. For example, link "strength" can be graded is a variety of different but somewhat overlapping ways. Links may vary in:

- Usage — how much traffic is observed over the link (e.g. packets processed)
- Capacity — how much traffic the link could sustain if pressed (e.g. bandwidth)
- Availability — how likely the link is to be available (e.g. how often a road is closed to traffic)
- Fuzzy membership, or "linkness" — how "link-like" is the connection between nodes (see below).

Each of these could be considered as representing "link strength" in some application. Incorporating any one of them in the description of a network will affect how the network should best be displayed in an application.

Links may vary in kind as well as in strength. For example, person A may send messages to person B, may dislike person B, may attend the same school as person B, may be stronger than person B, and so forth. Each of these possibilities is a relationship with the potential of defining a network among all the people under consideration. The same group of people may belong to a network of same or different school attendance, another network of message-sending and receiving, another network of liking and disliking, and so forth. Any one pair of nodes may be connected by a bundle of elementary links of different "flavour" or "colour", each of which may have different strength in one or more of the senses listed above. These bundles may be represented as links for some user tasks, and may need to be displayed split out into their elementary links for

other tasks.

The concept of a node, like that of a link, can be extended beyond that of the simple nexus where links meet. A node may process its input from one or more links before outputting traffic on other links. The output traffic could be of a completely different kind than the incoming traffic, and may have quite different temporal characteristics. A valve in a water-supply network, for example, has two inputs, one of water, and one of information (as to whether the valve should be open or closed). The information may arrive only as a momentary impulse that reverses the existing open-closed state, but the effect on the outgoing flow of water extends over indefinite time.

When there are many different kinds of links, as suggested in the above example of relations between person A and person B, the node itself might contain a network of such interrelations among traffic types. A node may therefore be even more complex than a link bundle.

2.3.1.1 Fuzzy nodes and links

The existence of a node need not be all-or-none. Imagine a road between two towns, A and B. The towns are nodes in a road network, and the road is a link in that network. Now a farmer builds a farmhouse near the road, at place X midway between the towns. Cars pass by without giving the farmhouse more than a cursory glance. Does place X constitute a node on the road network? Is the link between the two towns now broken into two links? Probably the answer is "No" to both questions.

Some years pass, and someone else has built a gas station and repair shop by the roadside near the farmhouse. Occasional cars stop there to refuel. Is place X now a node, and is the inter-town link between A and B broken into two? Again, probably not, but the answer is now less definite. For most traffic, the existence of the gas station at place X is unnoticed in the transit from town A to town B.

More years pass, and around the gas station and the old farmhouse, other people have built a pub, a furniture factory, and a few more houses. Occasionally travellers stop over for the night, and some people live and work there. Is the place now a node? Quite possibly it is, but it isn't a very good one, nor is the answer secure.

Finally, as more time passes, a few thousand people live around place X, the road is lined with shops, a church, a sports arena, and so forth, and the old farmhouse has been pulled down to make room for a new housing development. Clearly, place X must by now be a node on the road network, and what once was a single link between the two original towns now is a path over two links connected at the node that is the new town.

The farmhouse-to-town scenario illustrates that in a real world network, it is not always clear whether something is a node, and the same scenario can show that it may not be clear whether something is a link. This issue is different from evaluating the strength of an entity that is clearly a node (or a link). In the farmhouse scenario, the usage, capacity, and availability of the path from A to B remained unchanged, while it shifted from being one link to being a path over two links. The appropriate way to deal with this kind of situation is to say that the entity has a fuzzy membership in the class "node" (or "link"). For short, one might say it has a certain degree of "nodeness" or "linkness". Fuzzy membership of an entity in a class can range from zero to unity. In the scenario, before the farmhouse was built, the place had a membership of zero in the class "node", and even after the gas station joined the farmhouse, its "nodeness" was low.

When the membership of place X (the farmhouse region) in the class "node" was low, the membership of the

connection between towns A and B in the class “link” was high and those of the connections AX and BX were low. But as the nodeness of place X increased, so did the linkness of connections AX and BX”. At the same time, the linkness of connection AB decreased, until eventually it could not be considered a link. It had become a path over two links. All this changing of membership happened without any change in the ability of cars to go directly from town A to town B. What changed was the likelihood of cars stopping at place X.

All networks, even the most classical, can be treated as fuzzy, and sometimes it helps to do so. However, it is not clear how best to display fuzzy values. The problem is akin to that of displaying probability values or values that are uncertain. As mentioned above, there are at least four distinct concepts that can be considered as elements of link strength, traffic volume, traffic capacity, link availability, and “linkness”. All of these could simultaneously be of interest for any or all of the links in a network. The Framework cannot suggest how these should be represented ifor any particular task, but it can help the user or designer to recognize the possibility that they may need to be displayed.

2.3.2 Global dimensions of variation among networks

Extending the properties of individual nodes and links to allow for processing within nodes, multiple link types, and fuzzy membership are local to individual nodes or links of the network. Other, global, properties intrinsically concern more than one node or more than one link, and cannot be reduced to aggregates of local properties..

2.3.2.1 Striped networks

In the classical binary network, a node is a node is a node. If, however, the nodes of the network come in more than one flavour, it is possible that nodes of class A never are linked to other nodes of class A. In a purely heterosexual society, the mating relations among people would be represented by such a network. Males would mate only with females, and females only with males. Such a network is “striped”.

Striped networks are important in public health and probably in the spread of ideas as well. In either, a person at any particular moment may be designated as “infected’ (having accepted the idea), “immune” (unable to accept the idea), “susceptible” (not yet accepting the idea, but with a background into which the idea would fit), “carrier” (unable to accept the idea, but able to transmit it), and “dead” (which has no analogue in the transmission of ideas if one denies the Monty Python concept of the lethal joke).

“Stripiness” is a real-valued variable in a network, since rather than class A nodes never being linked to other nodes of class A, the mutual aversion may be only a statistical preference. Human mating relationships are not exclusively heterosexual, so the human mating network has a “stripiness” value less than unity. Stripiness can be considered a fuzzy property of a network.

2.3.2.2 Broadcast networks

A classical network might be called a “point-to-point” network. Each node is, perhaps fuzzily, connected to some set of other nodes. No matter how low the linkness of a connection, it is clear that traffic passes from a node only to those others with which it is linked. In principle, all potential receiving nodes could be found by tracing all the connections from that node that have a linkness greater than zero. A Broadcast network is different. Nodes that broadcast are not linked to any particular set of other nodes, and no search trace starting from an originating node could ever be sure of finding all potential receiving nodes. The traffic is available for reception by an unknown population of possible recipients, and that population might even have no members.

It makes no difference to the broadcasting node whether there are any or many receiving nodes.

A medical analogy might be the production of hormones or enzymes by some organ, which affect the operation of any systems in the volume bathed by their carrier fluid, including possibly the point-to-point interconnections of neurons performing quite different functions.

A point-to-point network is geometrically unidimensional; it consists conceptually of points connected by lines, no matter how complicated or fuzzy those interconnections might be. A broadcast network is at least two-dimensional, and may well be of much higher dimensionality. Of course, when displayed, the representation of a one-dimensional point-to-point network is usually embedded in a space of two or three dimensions, but that is an issue that we will deal with when we discuss embedding fields of networks and of displays.

Although broadcasting is a property of a single node, receipt of broadcast traffic is not. It depends on the possibility that there exists a population of nodes with the potential to receive the traffic. The "broadcast" property is therefore a network property, rather than a local property of the broadcasting node.

2.3.2.3 Stigmergic networks

A stigmergic network is one in which the node that originates the traffic simply emits it, leaving its effect to persist in the environment for any later suitable recipient or recipients to pick up. The ruts left in a muddy field, which tend to guide later vehicles, provide a simple example. That persistence makes it impossible for the links in the network to be precisely specified at the time the traffic is emitted. A stigmergic network could be a kind of broadcast network, in that the originating node is not, at the time the traffic originates, connected to any particular set of receiving nodes, and the traffic may be placed in a continuous field of any number of dimensions, not merely along prespecified lines of influence.

The difference between an ordinary broadcast network and a stigmergic one is that in a broadcast network the traffic can be received only at the moment it is available after transmission. If it is not received then, it is lost, whereas in a stigmergic network the effects of a traffic emission persist, and may affect nodes as yet unknown at times in an unknown future. The classic stigmergic example is the laying down of a pheromone trail by ants, which can be sensed by later ants to lead them to a source of food. If the ant had, instead, made a sound that could be heard by other ants, it would be broadcasting, and once it stopped making its sound, other ants would no longer be able to find the trail.

This temporal distinction between broadcast and stigmergic networks implies a fundamental difference in functionality between the two types. A broadcast network can sustain a conversation, in that if nodes have both transmitting and receiving capability, the receiver of an element of traffic (a message) can immediately broadcast another message that could possibly be received by the initial originator, thereby establishing a feedback loop between them. Of course, in a broadcast network, innumerable other nodes unknown to either partner may receive the transmissions of each. No such conversational feedback loop can be sustained in a stigmergic network, though stigmergic systems can show highly organized behaviour (as the ant example illustrates).

2.3.2.4 Traffic-free networks

Not all networks support quantifiable traffic across their links. In some, the structure itself takes the place of traffic. The network of family relationships is an example, in which only the actual links (brother-sister,

father-son) matter. Another, more serious, example is the network of influences among different parts of a city infrastructure. Such a network is implicit in risk assessment analyses such as what happens to which parts of the water supply if a particular electrical sub-station goes out of action, and what effect might that have on the food supply or the road traffic as refrigerators and traffic lights fail.

Networks of conceptual structures within and across documents are important for intelligence analyses, but no traffic flows over those conceptual links. The links may be directed, as in an implication, but directivity does not imply traffic flow.

Networks of resource conflict provide another example of networks without traffic. If each member of a set of teachers independently tries to schedule classrooms for lectures, it is probable that two or more teachers will schedule the same classroom for the same hour — a conflict. The whole set of links between teachers and classroom-hour requirements forms a network. Conflict occurs when a classroom-hour or a teacher is connected to two links, a satisfactory network in this case being one in which no path exists of length greater than unity. Analyses of such resource conflict networks could be useful in peacekeeping operations, if it can be shown that simple alterations of the network structure would eliminate the regions of conflict [5].

The discovery of the structure of some networks may depend entirely on observing the traffic over the (invisible) links, but this approach, with its related display implications, cannot be applied to the discovery and analysis of a traffic-free network.

2.3.2.5 Multidimensional description of network type in the Framework

Many networks combine point-to-point, broadcast, striped, and stigmergic properties in different combinations. News travels by conversation between individuals (point-to-point) as well as by TV broadcasts (broadcast) and by newspapers (stigmergic). The messages tramsitted by a radio broadcaster are received only by receivers, not by other transmitters (striped, broadcast). There properties are independent dimensions of description of networks, and the Framework should make it easy for the user to exploit that fact when looking for ways to display valuable features of the real-world network relevant to the task at hand.

2.4 Mathematical abstractions of network attributes

There is a large literature on the mathematical abstraction of various useful properties of classical binary networks and some of its extensions. Indeed, it sometimes seems as though many people think that taking one or more abstractions, such as one of the varieties of node centrality, or network diameter, or link density, or clustering tendency, and so forth, is sufficient to allow the user to understand the implications of the network. A Framework need not replicate the mathematical literature on networks, but it should attempt to categorize the different abstractions and use them to suggest appropriate display techniques as related to user tasks.

In the diagram of Figure 2, the Framework does not concern itself with the black box labelled “analysis tools”, but it is concerned with the set of network properties that the box can produce. The Framework group is beginning to catalogue different mathematical properties of networks, with a view to categorizing them in a way that will assist the user to select among display techniques that would suit the task at hand.

2.5 Network Dynamics

Traffic within a real-world network has a dynamical structure that depends on the properties of its nodes and links. If they are described accurately, along with the state at some epoch, then, in principle, the dynamical

behaviour of the equivalent mathematically abstracted network can be predicted. In the absence of influences from outside the network, the behaviour can settle into one of only three patterns: firstly a fixed point, in which traffic never changes over time; secondly a simple oscillation, in which traffic changes, but those changes repeat periodically; or thirdly, a "strange attractor", a pattern of changes that never repeats exactly, though it may appear to repeat approximately. The strange attractor is the signature of a chaotic system, which has behaviour that may well evolve quite differently from initial conditions that are very similar. A fixed point is the only possible terminal state of a network devoid of cycles.

Even very simple networks that contain cycles can have complex dynamical behaviour, so it is probably prudent in most cases just to assume that the behaviour is likely to be predictable only in the short term. Even if the network is truly isolated from external influences (and therefore quite likely to be uninteresting for real-world tasks), its dynamics are most probably chaotic, as chaos tends to be a property of feedback systems with nonlinearity as strong as a square law. If the nodes of a network do much transformation between their inputs and outputs, this condition is very likely to be satisfied.

On the other hand, adaptively evolved systems, including naturally grown networks, tend not to behave fully chaotically, but to reach a state often known as "on the edge of chaos". Such systems are usually robust against disturbances and failures, but occasionally fail dramatically and unpredictably. They can in many cases be strengthened by subjecting them to frequent small disturbances and failures, to allow them to adapt to an error-tolerant condition.

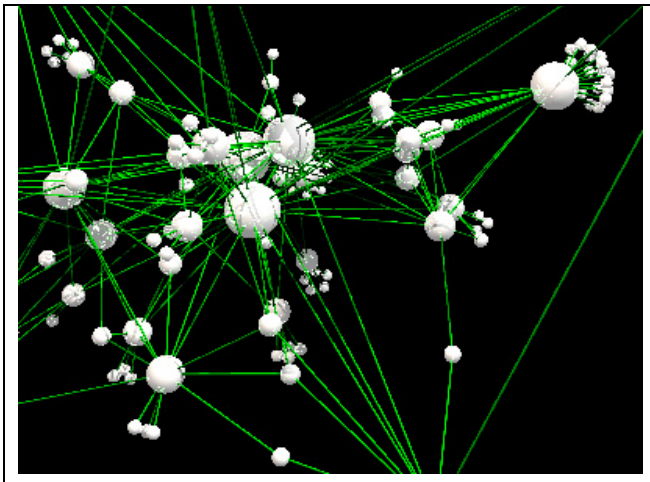
Displays of network dynamics are difficult to incorporate within the Framework, but a good Framework should at least suggest dynamical attributes that might be worth representing in the display.

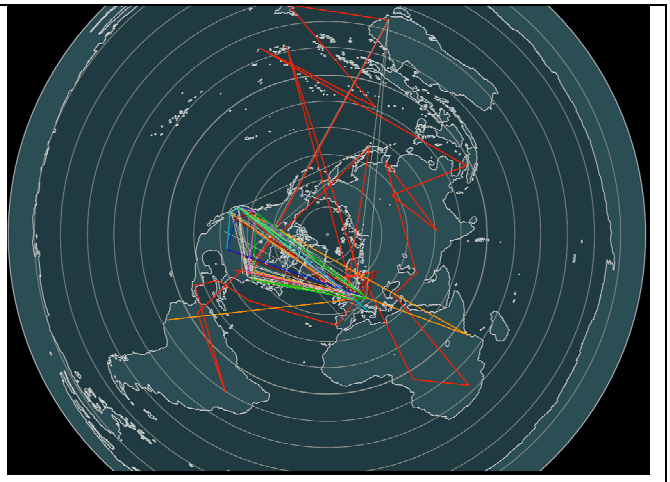
2.6 Embedding Fields

A network in the real world consists of physical entities connected by relationships that may be physically embodied (e.g. roads, wires) or purely conceptual (family tree, social influence, etc.). The network is embedded in a physical or conceptual substrate, but what determines its "embedding field" is the set of contextual attributes in which changes make a difference to the network from the viewpoint of the user and for the user's current purpose.

The embedding field can be thought of as the currently relevant context. For example, a road network exists in a landscape of hills, valleys, rivers, and towns. It may make no difference to the traveller where the road is laid between towns, provided it is not too winding or steep, but it does make a difference to the people who live and work near the roads. For the traveller uninterested in the view, the embedding field may consist simply of the choice points and travel distances; for the local inhabitant, it is the geographical landscape.

For purposes of display, the embedding field can provide context that may suggest to the user the real-world meaning of network properties being displayed. Figure 5 shows two examples. In one, no embedding field is shown. Spheres represent Web pages, and the distance between two spheres represents the similarity among the topics of the two pages. In the other picture, the interesting question is how places in the world tend to be linked by Web traffic on a particular topic. The world map provides the context that makes the data intelligible.





5. Two views on parts of the World Wide Web. The left picture (5a) shows topical relationships, the right one (5b) traffic in a geographic context (an embedding field for the network) (Images are from http://www.visualcomplexity.com/vc/, with permission of the respective authors).

2.7 User tasks

User tasks, like the representable properties of networks, should be categorized in any reasonable framework for network visualisation. One clear classification is into tasks that start with the user analysing aspects of the structure or behaviour of a network for which the actual nodes and links are known, as opposed to tasks that require the user to discover the existence or the connectivity of important parts of its structure.

2.7.1 Network discovery tasks

If the user wants to discover the structure of a network, the perceptual mode must almost always be Exploring. Sometimes in may be Search, if the need is to discover some implicit but as yet unknown part of the network for some immediately current purpose, such as discovering the recent social contacts of a person who has just come under suspicion. But usually the objective is to discover something about the network structure in order that other information can be incorporated at some later time. The network itself would be presumed to have a structure stable enough that the structure discovered would remain useful for those later purposes.

The evidence that enables network discovery is ordinarily the discovery of a link connecting a known node to somewhere else, The existence of the link implies the existence of a node at its other end, which may not have been known before the discovery of the link. Link detection often is possible only from the observation of traffic, so the display issue is likely to involve the integration of observed traffic, to form the links and thereby identify many of the nodes in the network. However, if the network is traffic-free or stigmergic, there is no traffic to follow. Other techniques, implying other kinds of display, must be used. For example, in a network of conceptual relations, something about the characteristics of the nodes themselves may allow nodes that should be linked to be correlated. Network discovery might then be aided by display of the correlations of their attributes among different nodes.

In a stigmergic network, links may sometimes be discovered and refined by exploring the embedding field rather than by exploration within the network itself. The classic example of a stigmergic network is provided by the ant pheromone trail. The trail is a one-dimensional link between the ant nest and the food source. It emits pheromones into the air, a 3-dimensional space. A sniffer could detect the existence of a pheromone link. It might then be able to refine its detection to the fact that the pheromone source was ground-based (a 2-D embedding field), and by scanning across the ground, might be able to trace the one-dimensional path of the trail. Even though a stigmergic network cannot be discovered by analyzing traffic patterns, the potential for one to exist might be found by tracing the persistent effects that might be picked up by later agents that would become nodes in the stigmergic network.

2.7.2 Network Analysis tasks

Network analysis is not always functionally distinct from network discovery, though it is always conceptually distinct. Functionally, analysis of part of a network may lead to suspicions that other parts of the network remain to be discovered, and the attributes of parts that are newly discovered may then need to be analyzed. Conceptually, though, analysis is conducted only on parts of the network that have already been discovered.

Analysis tasks may be categorized, like network properties, into those that look for local properties such as the most central node (in some definition of centrality), and those that compute properties of the network as a whole, such as its diameter, its structure (e.g. random, scale-free, small-world, etc.), its cyclicity, and so forth.

Analysis tasks can be categorized in a different dimension: description versus construction. A descriptive analysis has the result that the network's properties or behaviour is in some way represented in a form that can be displayed. A constructive analysis is done in order to modify the network to make it have some desired property such as robustness against loss of some nodes or links. Descriptive analysis for later use implies the Exploring mode of perception, whereas descriptive analysis to seek out significant structural or dynamic patterns would be Search, perhaps augmented by some Alerting performed by autonomous agents that apply algorithms to parts of the network and signal when their algorithm provides an appropriate output. Constructive analysis uses the Controlling mode of perception, continuously observing the critical properties as changes are made in the network. It also may be supported by Alerting, if autonomous agents are set to warn of undesirable behaviours they had been programmed to recognize.

Descriptive and constructive analysis are likely to benefit from different kinds of display, the descriptive probably requiring more display of context, the constructive perhaps involving more direct user interaction. Obviously, these differences can only be tendencies, but the Framework should assist the user to recognize whether they apply in the particular task.

2.7.3 A preliminary taxonomy of user tasks

From the discussion above, user tasks may be categorized as:

- Network discovery (probably using the Explore mode of Perception)
  - Link-based
  - Node attribute based
- Network Analysis (any mode of perception)
  - Local
  - Global
    - Descriptive

- Constructive

Each of these categories is likely to have implications for the kind of display most suited to the task. Of course, within these gross categories there will be many refinements and idiosyncratic requirements for displays in different domain contexts. In the sense of the RM-Vis Reference Model, this taxonomy is only one dimension of the description of user tasks.

2.8 Data and Display typology

The Final Report of IST-013/RTG-002 (*Visualisation of Massive Military Data Sets: Human Factors, Applications, and Technologies*, NATO RTO-TR-030, known as the HAT Report) provided a taxonomy of data types, reproduced in Table 1 (from Table 3.1 of the HAT report).

This taxonomy identifies six characteristics of data, each of which may take on different kinds of value. The particular values for different attributes often affects the choice of ways to display the data. For example, if the data has symbolic values, it will probably be best displayed in some way involving symbols, whereas if its values are analogue, graphs or continuum displays such as grey-scale, symbol size, or colur value may be more appropriate. Likewise, streamed data often implies a different kind of display than static data, if only because streaming carries the implication that a significant event may occur at any moment, and the user should be able to detect the occurrence of the event when it happens. In other words. Streaming data implies the probable use of the Alerting Mode of perception, whereas static data ordinarily does not, though it may.

The HAT report typology considered data in any domain. Some refinement may be possible when the domain is restricted to tasks involving networks. In many such tasks, the structure of the network is likely to be static within the time frame of the analysis. Traffic on the network, however, is more likely to be streamed, which suggests that a different kind of display might be suited for analyzing network structure as opposed to network traffic. On the other hand, the task of network discovery may well involve the continuing addition (and perhaps deletion) of links and link types. In that case, Alerting might well be used even though the data themselves are static.

The HAT Report also considered a taxonomy of display presentation types, shown in Table 2 (Table 3.2 of the HAT Report). In this table the word “Data” refers not to the incoming application data, but to the data actually to be presented to the user. It may be a static picture or may vary dynamically; the user may control shich data are selected for display, or they may be selected according to some algorithm over which the user has no control; the data may be identified by some label, or may intrinsically be identified only by a location, as is usually the case for georeferenced map data; and finally, the data values may be analogue or categoric, and if categoric may have syntactic interrelations like the words of a language, or may be isolated representatives of labelled classes. All these characteristics could potentially vary independently. They interact with the application data characteristics, but are not the same thing. Network connectivity, for example, can be represented by a matrix (categoric non-linguistic) or by a pattern of lines (analogue vector).

Any kind of application data (from Table 1) could, in principle, map to any type of display representation. Some of the application data types, however, map naturally onto the display types. For example, symbolic linguistic data may well be displayed in a categoric linguistic manner, such as text or tables. The HAT Report listed a few of these natural mappings, as shown in Table 3 (Table 3.3 of the HAT Report).

Table 1. Summary of Data Types

Acquisition	Streamed	regular	
		sporadic	
	Static		
Sources	Single		
	Multiple		
Choice	User-selected		
	Externally imposed		
Identification	Located		
	Labelled		
Values	Analogue	scalar	
		vector	
	Categoric (classic)	symbolic	linguistic
			non-linguistic
		non-symbolic	linguistic
			non-linguistic
	Categoric (fuzzy)	symbolic	(non-linguistic)
		non-symbolic	(non-linguistic)
Interrelations	User-structured		
	Source-structured		

Table 2. Summary of Display Presentation Types

Display Timing	static	
	dynamic	
Data Selection	User-selected	
	Algorithmically directed	
Data Placement	Located	
	Labelled	
Data Values	Analogue	scalar
		vector
	Categoric	linguistic
		non-linguistic

Table 3. Some examples of mapping data types onto display types

Data type	Display type	Comment
Streamed	**Dynamic**	The user ordinarily wants to act when some event occurs
Located 2-D or 3-D	**Located**	The display is a 2-D or 3-D map of some attribute(s) of the data. If the location identification of the data is in a higher dimensional space, there is no such natural mapping. Tricks must be used.
Labelled	**Labelled**	The display is likely to be tabular, or some kind of a graph such as a histogram or pie chart
Analogue scalar	**Analogue scalar**	Even if the data are identified by label, its analogue values map naturally to analogue display variables such as the length of a line or the brightness of a pixel.
Analogue vector	**If 2-D or 3-D, Analogue vector**	A 2-D attribute can be mapped onto an area display, a line with length and orientation, a colour hue, a sound location, a sound intensity and pitch, and so forth, all analogue vector attributes of the display. A 3-D attribute can similarly be mapped into a volumetric display. Higher dimensional analogue attributes can be displayed, but the mapping is less obviously "natural."
Categoric	**Categoric**	Categoric data values have no natural relation to analogue display values, and must be displayed categorically. The categoric display attributes may or may not map "naturally" onto the categoric data attributes. This kind of mapping is usually considered to be "cognitive metaphor."

2.8.1 Embedding fields of displays

We have discussed the concept of the embedding field of a network, and whether or when a network's embedding field should be displayed. Displays themselves have a hierarchy of embedding fields, which perhaps are worth considering, especially when dealing with the display of the embedding field of a network.

Consider a display on the screen of a normal computer monitor. It consists of, say, 1600 x 1200 or 1,920,000 different pixels, each capable of showing, say, 256 different intensities of red, blue, and green, and of changing each of these three values 75 times per second. That amounts to a potential information bandwidth of 3,456,000,000 bits/second. Clearly, humans can process only a very tiny fraction of that, so actual displays for human consumption must be highly redundant. It is this redundancy, expressed hierarchically, that allows us to talk about embedding fields of display.

A redundancy giving rise to the first level of embedding might be correlation between neighbour pixels in time and space. If a pixel has certain values of red, blue, and green at one moment, it is very likely to have similar values of red, blue, and green 1/75 second later, and all the pixels in its neighbourhood are likely to have similar values. Only at very few places on the screen and moments in time will neighbouring pixels have

distinctly different values of red, blue and green, and if a pixel is in such a place, other nearby pixels probably are, too. Those places are called “edges” in a picture. Most pictures, including text, consist largely of more or less uniform areas that persist for substantial times, and which are separated from each other by relatively sharp changes. This coherence allows a second level of embedding, of visual objects constructed from coherences among the persistent patches of colour.

A display of a network could use the level of visual objects as its embedding field, some objects representing nodes, some representing links, and the rest forming a background field. Such a display would be too plain for any but the simplest of classical networks. Further levels of display embedding offer richer possibilities. Objects may be coordinated, using more redundancy but making the display more intelligible. Different kinds of coordination might give the viewer the impression of looking into a 3-D space with objects grouped at different depths. Such an impression can be developed through coordinated movement patterns, simulated overlap of the objects represented, use of colour to simulate atmospheric effects, and so forth.

A 3-D field can in this way be embedded in a 2-D field, but only at the cost of reducing the potential information capacity of the display. The benefit is that the human would be unlikely to be able to use the available 2-D capacity, so the redundancy inherent in the embedding process does not necessarily reduce the information transmitted to the human viewer. Indeed, it is likely to increase the transmitted information, because of the improved interpretability of the resulting display — its increased likeness to what the viewer sees in everyday life.

There are many ways of using redundancy to trade the excess information capacity of the display for increased intelligibility. Each stage of increased redundancy provides an embedding field of display in which the next, more redundant, display domain is shown; the field of patches is embedded in the pixel field, the objects in the field of patches, the 3-D space in the 2-D field of objects, and so forth. When it comes to displaying some aspect of a network that is the focus of the viewer’s interest, it is tempting to suggest that the hierarchy of embedding fields of the display might be used to show the embedding field (the context that supplies the meaning) of the network. Indeed, this is what is done in Figure 5b, but not in 5a.

Information-theoretic approaches to display generally [6], the display of networks [7,8], and to the analysis of redundancy in embedding have not yet been integrated. Such integration may well prove useful in developing the Framework for visualising networks beyond what is anticipated within the life of this Working group. The notion of “information” carries with it the concept of uncertainty, and the display of uncertainty, as well as of fuzzy membership, is another area in which display redundancy must be used.

3 FRAMEWORK AS PROCESS

The VisTG Reference model implies a process for developing displays that assist visualisation. It starts with the question of what the user wants to see, and progresses through a codified series of six questions and answers relating to how that desire can be satisfied. Some of the answers imply that the user might want to see something else, and each such requirement leads to its own set of six questions. For example, if the user wants to see the part of the network within two steps from the node having the highest centrality value, there is an implication that he will want to see the controls that tell the box of analysis tools to find the node of greatest centrality, and that he will also want to see the controls that prune the display according to distance from a selected node.

It is not necessary to go into detail of the implications of the VisTG Reference Model, which can become

quite complex despite the apparent simplicity of the primary diagram (Fig 3). However, when it is known a priori that the visualisation task relates to a network, it becomes feasible to begin to categorize what the user might want to see. The possibility of such a categorization leads to a different process sequence, based on investigating which aspects of the network are relevant to the user's real task, and then inquiring how best to display them.

Imagine a simple scenario. Some people have been placed under surveillance for what appears to be good legal reasons. Some of their social contacts have been observed, the records of the addresses of their e-mail contacts have been recorded, as have the numbers of incoming and outgoing calls on their telephones. There is a large, but incomplete, set of data that can be construed as a network. How would an adequate framework help, say, an analyst to interpret possible threat (or to dismiss that probability)?

According to the VisTG Reference Model, the first step is to ask what the user wants to do. In the network context, the preliminary task taxonomy (Section 2.7.3) asks whether the user is interested in network discovery or analysis of a known network. The data do not supply the answer. In the scenario, the answer may depend on whether the authorities anticipate being able to test for contacts of different types when the analyst asks, or whether more data can be supplied in respect of people outside the original group for whom legal authority for surveillance had been approved. On the other hand, it might depend on whether the sheer quantity of data available to the analyst made the extraction of interesting subnets a major part of the problem. It all depends on the analyst.

Let us suppose that the analyst decided that the problem of the moment is network analysis, to determine whether individuals in the group, or the group as a whole, have any suspicious pattern of activity. The next thing is for the analyst to determine what characteristics of the network might bear on the answer. The Framework categorization of network properties should assist in this determination, but the Framework cannot be expected to provide any definitive mapping between an open-ended set of possibly interesting patterns and a finite but large set of computable network properties.

Having settled on a set of probably useful network attributes, the analyst needs to find out how best to display them, and whether applications exist that can provide such a display. Here, the Framework should be able to help, through its categorization of network attributes, both local and global. The four-dimensional characterization of networks and their attributes (Section 2.3.2.5) should be valuable for this part of the process. For example, in the scenario, it is probable that the public roles of the social contacts of the group under surveillance will be known, which means that in the network, they could be designated as different types of nodes, which implies that attributes related to stripiness might be useful.

Once the attributes and the desired display techniques have been specified, the analyst's next task is to discover what tools are available to extract those attributes and produce those displays. It is here that the Framework should be linked intimately with the IST-059/RTG-025 Survey of network visualisation applications. The user's requirements that are characterized by the Framework should now be used as a filter to select from the Survey Database those applications that can satisfy them. From the filtered list of applications, the analyst should be able to select the most appropriate, knowing what to expect and how to get the answers to the real-world task.

4 THE WAY AHEAD

At this stage, the Framework is little more than a set of concepts. There are many stages before a full

Framework can be considered complete. For one thing, the conceptual categorization of network attributes should be applied to the attributes that are ordinarily considered in Social Network Analysis, and gaps in the available models should be made explicit. For example, as compared to what is available for crisp point-to-point networks, there seems to be not much available for broadcast, stigmergic, or fuzzy networks, which are a large part of the real world of networks.

The first steps toward a mutually consistent categorisation of tasks, display techniques, network attributes have been taken, but the implied mappings have barely been considered. The information-theoretic implications of networks and displays should be useful in suggesting appropriate mappings, but is only one possible approach among many possibilities.

Mapping is an integral aspect of the Framework. For which kinds of tasks should which representable attributes be selected, and how should they best be displayed for which kinds of user? Little work has been done on this beyond what was reported in the HAT Report [1] and reproduced above (Section 2.8).

During the development of the categories and their mappings, the Framework should be tested against scenarios from different application domains, as well as against the multitude of display technologies and software packages available on the World Wide Web and elsewhere. One problem with doing so is that the tests then become biased in favour of what is already known, whereas the Framework ought also to highlight areas in which new research or development could be both useful and cost-effective.

In parallel with much of the above, software should be developed to provide a supportive environment within which the Framework could be applied to user tasks. This software would implement the "Framework as Process", whereas the foregoing would implement the "Framework as Structure".

When the Framework seems to have become moderately useful and useable, it should be privately demonstrated to representative users, such as both high and lower-ranking military personnel, and selected government researchers and developers, perhaps at a workshop such as the present one. Comments should then lead to an iteration in the development of the Framework.

The final step in the development of the Framework is its release in "alpha" form to the general public. Thereafter, its further development would be a matter of the public record.

REFERENCES

[1] NATO Technical Report RTO-TR-030, *Visualisation of Massive Military Datasets: Human Factors, Applications, and Technologies,* 2001.

[2] Gouin, D., Vernik, R. J., Evdokiou, P., Houghton, P., Barnes, M., Monk, D., and Rosenblum, L.. *Information Visualization in C3I: Final Report of TTCP C3I AG-3*. DOC-C3I-AG3-1-2002. TTCP C3I AG-3, 2002.

[3] Taylor, M. M., Perceptual principles relating to remote sensing. *Proceedings of the First Canadian Remote Sensing Symposium*, 1972, Vol 2, 497-503

[4] Cunningham, W. B., and M. M. Taylor, Information for Battle Command, *Military Review*, 1994, Vol. 74 no. 11, 81-84.

[5] Kearns, M., Suri, S., and Montfort, M., An Experimental Study of the Coloring Problem on Human Subject Networks, *Science*, 313, 11 August 2006, 824–827

[6] Smestad, T., Outline for concepts and principles of effective visualization, Annex A1 to the *Final Report of IST-021/RTG-007*, 2004

[7] Bjørke, J.-T., Map Generalization of Road Networks. In *Proceedings of IST-043/RWS-006, Visualisation and the Common Operating Picture*, 2004.

[8] Bjørke, J. T. and E. Isaksen. Map Generalization of Road Networks: Case study from Norwegian small scale maps. *Proceedings XXII International cartographic Conference*, La Coruña, Spain, 11-16 July 2005.

A Framework for Network Visualisation

Progress Report

Framework Working Group of IST-059/RTG-025

(J.-T. Bjørke, M. R. Nixon, M. M. Taylor, A. K.C.S. Vanderbilt, M. Varga)

A Framework for Network Visualisation

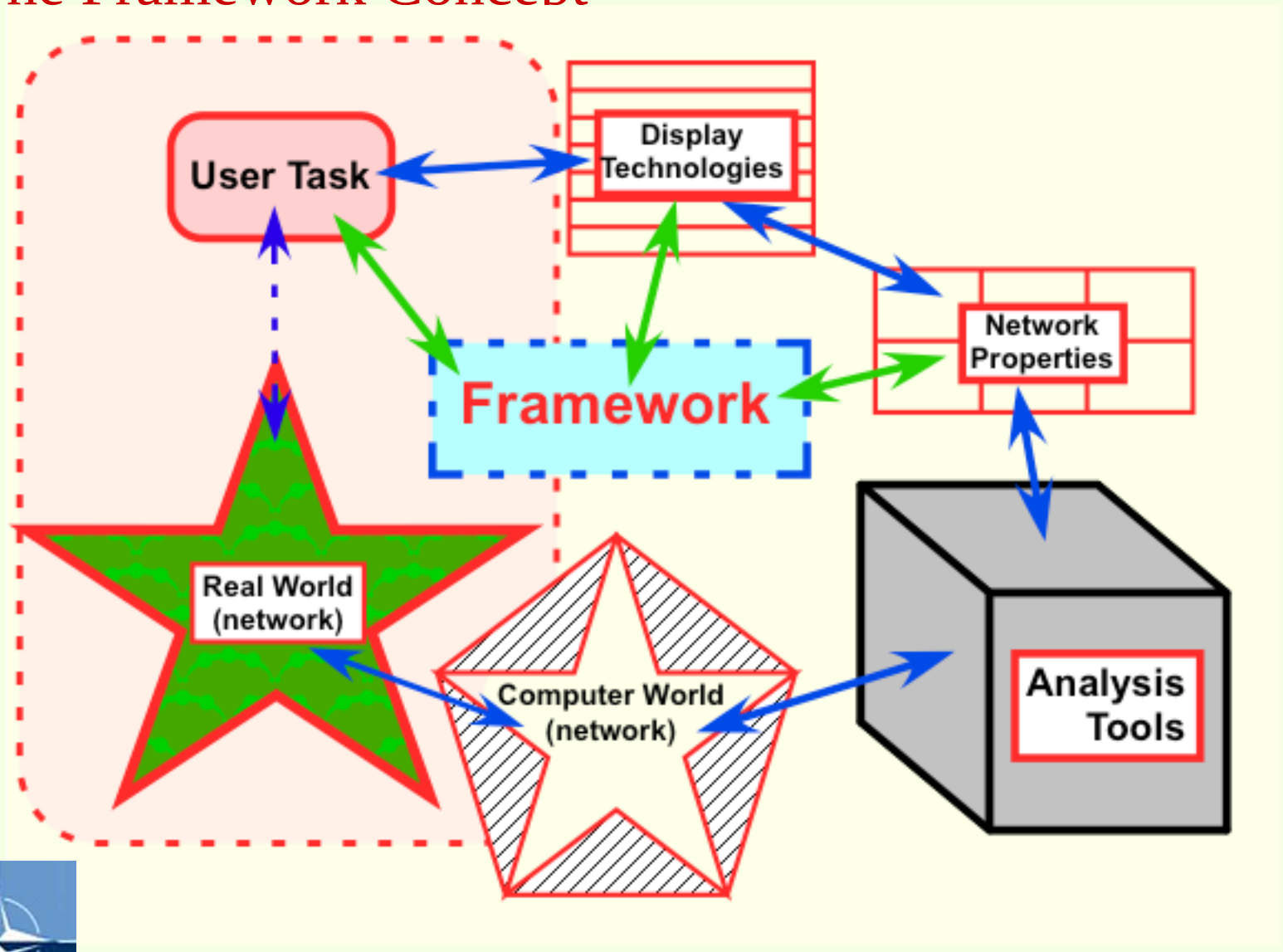
The User's Problem: How to coordinate

- The user wants to solve some real-world difficulty that in some way involves a network.
- The real-world data are abstracted into computer data that can be construed as a network.
- Algorithms abstract both local and global properties of the network that might be useful for the user's real-world task.
- Properties of the network likely to be useful are displayed.
- The display helps the user to visualise the state of the real world in which the difficulty exists.

Framework

- A Framework for network visualisation should tie together these elements in a coherent way, relating task to display, and display to network properties.

The Framework Concept



What is a “Framework”?

The Working Group conceives a Framework for Network Visualisation as

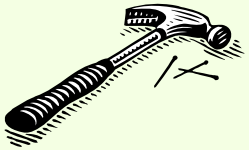
- An interface that connects a network-related task requirement with the available display technologies
- An interface that connects the available display technologies with computed network properties
- A way of categorizing and describing user needs, display technologies, and network properties
- A help to users in assessing the nature of their requirements
- A guide to users in choosing a visualisation system suitable for their application need.

- A guide to developers and researchers as to unmet needs.

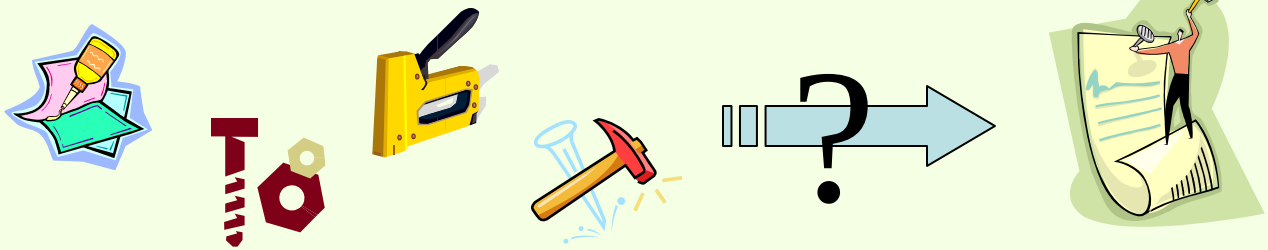
Why is a “Framework”?

The question may be ungrammatical, but it is significant.

Why is IST-059/RTG-025 concerned with the task of developing a framework in the first place?



- *If I have only a hammer, every job seems to require nails.*
- *If I need to fasten something, how do I know hammers exist?*
- *If I need something fastened and I know the tools exist, do I glue, screw, staple, or nail?*
- “I” would want a Framework that categorized fastening jobs in terms of what tools were best for those jobs, and categorized tools in terms of what kinds of fastening jobs they did best.



Why a Framework for ***network*** visualisation?

- Numerous ad-hoc examples of network representations have been created for specific applications, some of them very good for their purpose.
- It is usually not clear how the insights that led to particularly effective representations can be generalized to new situations.
- A good Framework should help identify the conditions for which different insights are helpful.
- Users need to see different aspects of network structure and function, and some of those aspects are not well served by extant display techniques.
- Users usually choose to see those aspects for which effective display techniques are available (they are given only a hammer!).
- A good Framework may help inspire research on new modes of display for different kinds of network properties.

The form of a Framework for Network Visualisation

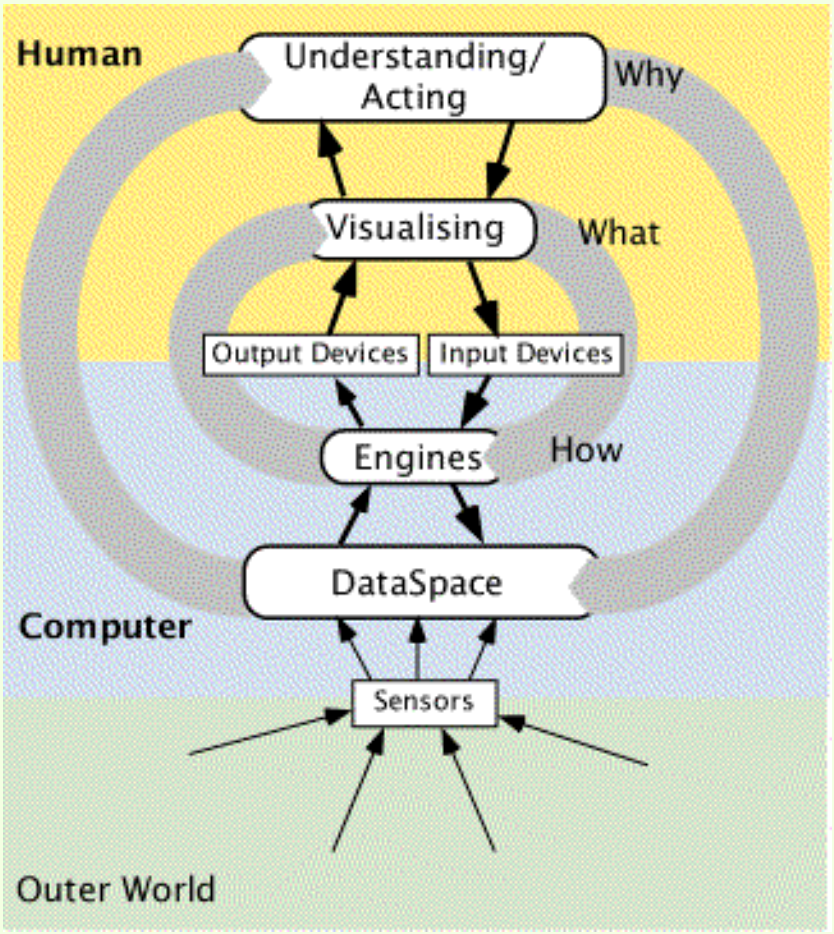
A Framework for network visualisation should include:

- A structured approach to describing user needs
- A structured set of displayable properties of networks
- A structured way of describing display techniques
- A process to help the user match needs to displayable properties using the appropriate display techniques.

Framework roots: Visualisation Reference Models

- VisTG Reference Model
  - *A functional model developed initially by predecessor groups of IST-059/RTG-025*
  - *User's purposes determine the representation characteristics*
  - *Separate interaction loop levels for primary tasks, algorithms and engines, and interface*
- RM-Vis Reference Model
  - *A descriptive model developed initially by a working group of The Technical Co-operation Programme (TTCP) C3I AG-3*
  - *Separable dimensions of description for application domain, content to be displayed, and display approaches*

The VisTG Reference Model



The VisTG Reference Model has 3 loops, the outer acting through the inner:

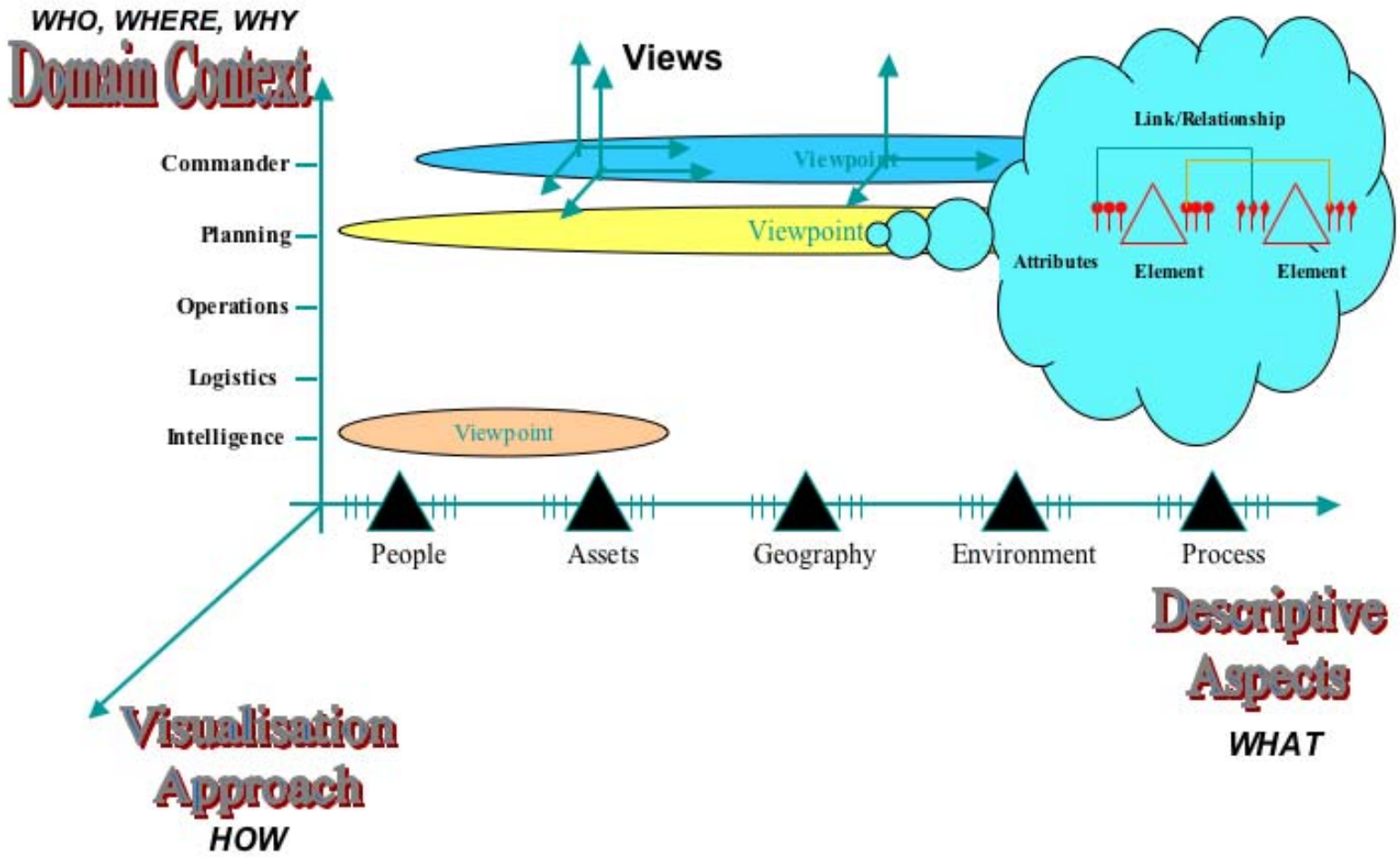
(1) The user understanding and acting on the data in the dataspace, which involves...

(2) The user visualising the data provided by and massaged by the engines under the control of the user, using...

(3) The Input-Output devices that interact with the user's sensors and musculature.

But we assume that the user "really" wants to influence the outer world!

RM-Vis Framework (Overview)



RM-Vis Reference Model developed by TTCP C31 AG-3

Report to IST-063/RWS-010 by the IST-059/RTG-025 Working Group on Framework for Network Visualisation

Framework: Categorizing user tasks

It has proved useful in the past to consider four distinct modes of perception. These suggest approaches to information display, and are equally applicable to guide categorization of user tasks.

Perceptual Modes

1. Controlling/Monitoring: *Keeping track of a changing situation and possibly acting to alter it.*
2. Searching: *Looking for something immediately wanted*
3. Exploring: *Building understanding of slowly varying context that could be useful for later search or control.*
4. Alerting: *Marking that a prespecified condition has occurred in a datastream or exists within a large dataspace. Alerting is usually an automated process.*

Framework: Perceptual modes

One of the dimensions of the RM-Vis reference model is “Domain Context”, which specifies an application area. Each domain context has its own specific possibilities for the four perceptual modes, so the Framework does no more than to suggest to the user that the requirements be identified in each of the four modes.

For example, in an anti-terrorist application,

Exploring might use network analysis to identify groups of people worth

Monitoring, while

Searching might seek those among them with contacts in specific areas of expertise, and

Alerting might set up automated procedures to look for certain types of traffic in particular areas of the identified network.

Each of these implies different requirements for display.

Framework: Perceptual mode implications for display

The four modes in the anti-terrorist scenario:

1. Exploring involves the discovery of networks, and might benefit from a fisheye display of the portions of the network so far discovered.
2. Monitoring implies continuing analysis of traffic dynamics, and requires the ability to dive into detail at specific moments.
3. Searching concerns the attributes of specific nodes, to discover their potentialities when matched with those of linked nodes, and hence requires both wide range and closely focused display representations.
4. Alerting is a programmed background activity that suggests the requirement to display relevant aspects of the network in context, when any of the prespecified patterns is detected.

Framework roots: Network properties

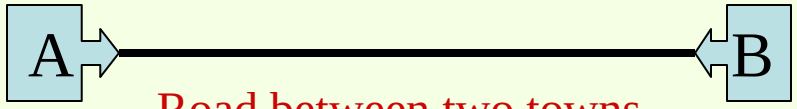
- Network types:
  - *Point-to-point, broadcast, striped, stigmergic, fuzzy or crisp*
- Mathematical relations and functions in abstract networks
  - *Many important representable properties*
- Dynamic properties of real networks
- Transformational properties of nodes and links of real networks
  - *Inputs may be of different nature to ouputs*
- Embedding fields of real networks and of displays
  - *Determine and constrain potentialities of the network*
- Data Source: static or streaming, and other properties
  - *Is the network changing while the user watches?*

Network Types

- Point to point *The classic network. Nodes are defined and each node is or is not linked to each other node.*
- Broadcast *A transmitting node cannot know which of many eligible receiving nodes may receive the traffic (e.g. airborne infection).*
- Stigmergic *The "traffic" is left in the environment and may be received at an indeterminate later time by an indeterminate number of receivers (e.g. ruts that tend to guide later traffic, etc.)*
- Fuzzy *An entity (node or link) is not well defined. Nodes may be somewhat linked to other nodes (e.g. suitability of road for heavy traffic). The degree of linkage may depend on the user's purpose.*
- Striped *Nodes of type A can be linked only to nodes of type B and vice-versa (e.g. vector-transmitted diseases).*

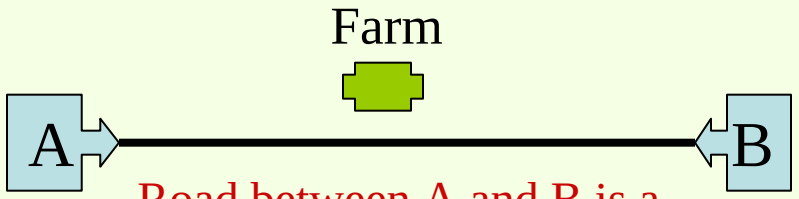
Fuzzy Nodes and Links — simple example

1 Original situation



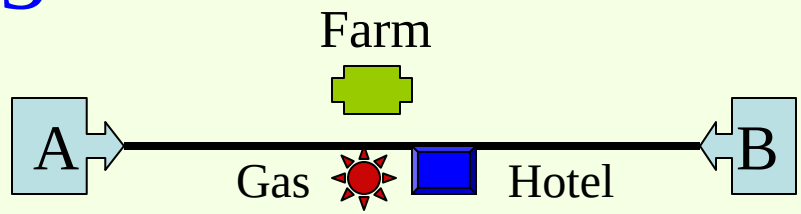
Road between two towns A and B is a link

2 A Farmhouse is built



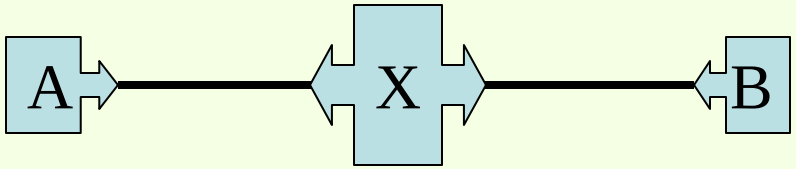
Road between A and B is a link? Yes, Pretty much. Is the farm a node? Not really.

3 More facilities are built



Road between A and B is a link? Sort of! Is the building cluster a node?

4 The cluster becomes a new town



Road between A and B is NOT a link. Roads A-X and B-X are links, and X is clearly a node by now.

Varieties of Link "Strength"

In many displays of networks, "strong" links are shown more vividly than are "weak" links. However, links have several independent parameters that might be called "strength"

- Utilization — if the link is of a kind that has traffic, how much is there?
- Capacity — How much traffic could the link sustain?
- Availability — What is the probability the link will be open for traffic?
- Coherence — (Of a traffic-free link)

  How tight is the relationship between the terminal nodes? (*sibling is tighter than second cousin*)

- Fuzzy membership — How much like a link is the connection?

How should these characteristics be distinguished in displays?

Varieties of Link "Strength"

A link may be simple, carrying one kind of traffic or representing one relationship, but what seems to be a single link might actually be a bundle of elementary links of different kinds.

For example, person A might at the same time

- be the father of person B,
- lend money to B,
- enjoy B's company,
- telephone B frequently.

How should a "bundle" link be distinguished in displays? Is the number of elementary links another dimension of link "strength"?

The complexity of a link bundle implies that the nodes it links are themselves complex, perhaps including a whole network that interconnects the elementary links of the bundle. *How should that complexity be displayed?*

Transformational properties of nodes and links

In an abstract mathematical network, a node may be only a place where traffic enters and is distributed to outgoing links.

In a real network, the nature of the traffic and its timing are determined by processes that occur in the node.

Example: a person (a node) may receive messages from a variety of sources over a period of time, may interpret the messages, and may take action that affects other people, but not by sending messages.

- *Point-to-point gossip about the evil effects of immunization may cause a parent not to immunize a child, who then catches and propagates a serious disease;*
- *Alternatively, broadcast messages may induce sufficient people to get immunized that a potential pandemic is avoided.*

The network in this example contains both broadcast and point-to-point elements, the links are fuzzy, and the nodes significantly transform their inputs in generating their outputs.

Mathematical Properties

Most of the mathematical properties have been developed in connection with crisp point-to-point networks. Some examples:

- Network connectivity: *random*, *scale-free*, *tree*.
- Centrality: *distribution of linkage degree over the nodes*
- Directivity: *Whether links are unidirectional or two-way*
- Cyclicity: *Can traffic go from A through other nodes and back to A?*
- Diameter: *The longest geodesic between any pair of nodes*

The mathematical properties of fuzzy networks should reduce to those of crisp networks in the limit of binary membership functions (only zero or unity allowed), but are less well developed.

Real Networks

Are not mathematical abstractions.

They are messy.

They are embedded in a complicated environment

They are not well-defined or completely known

They **are** what real users have to deal with.

Embedding fields of real networks

Assertions by Joanne Treurniet (IST-059/RTG-025 member):

1. A physical network always has the possibility that a conceptual network lies on top of it. The conceptual network may map homologously onto the physical network if the relationships between nodes are defined as such, but in most cases, the conceptual network involves only subsets of the physical network.

2. A conceptual network may exist without any underlying physical network.

Examining these assertions led to the concept of an "embedding field" for a network.

Embedding fields of real networks – 2

- A network in the real world consists of physical entities connected by relationships that may be
  - physically embodied (e.g. roads, wires) or
  - purely conceptual (family tree, social influence, etc.)

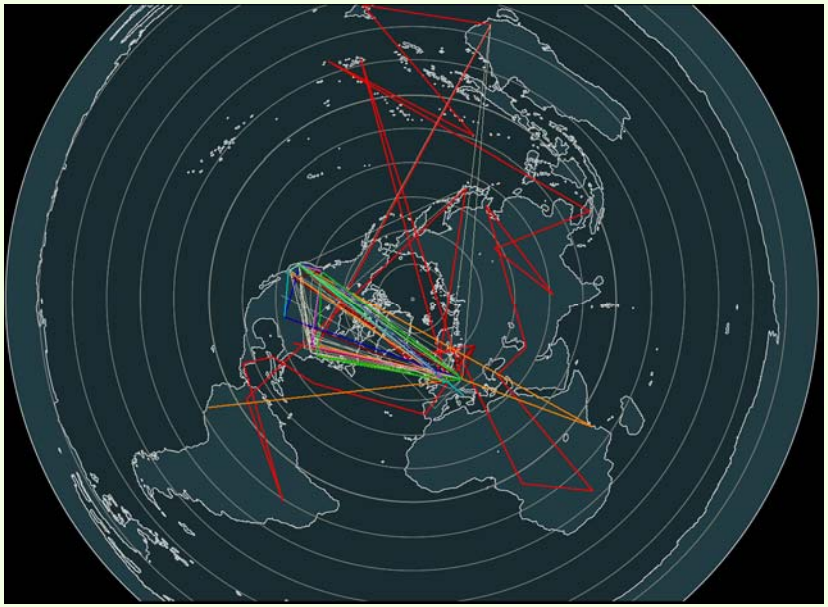
- The network is *embedded* in a physical or conceptual substrate, but what determines its "embedding field" is the set of contextual attributes in which changes make a difference to the network ***from the viewpoint of the user and for the user's current purpose.*** The embedding field can be thought of as the currently relevant context.

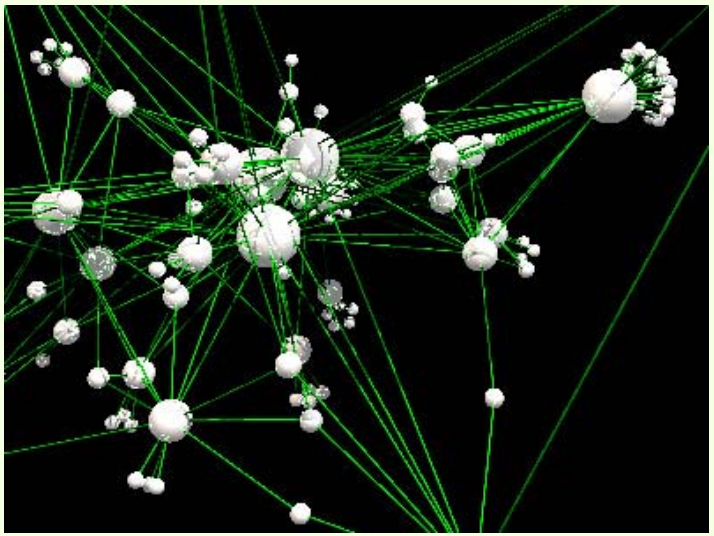
- For example, A road network exists in a landscape of hills, valleys, rivers, and towns. It may make no difference to the traveller where the road is laid between towns, but it does make a difference to the people who live and work near the roads. For the traveller uninterested in the view, the embedding field may consist simply of the choice points and travel distances; for the local inhabitant, it is the geographical landscape.

Embedding fields of real networks – 3

Networks are often displayed along with some aspect of their embedding field to supply context.

But not always:





Two representations of part of the Internet.

Embedding fields of real networks – 4

The embedding field for a network is often another network

- *e.g. for a contagious disease, the network of infections is embedded in the network of social contacts, but for an airborne disease it is not.*

Networks can inherit properties from their embedding fields

- *e.g. location for a geographic embedding field, contacts for a social relationship network embedding field.*

The embedding field constrains the properties of the embedded network, but new attributes can be developed

- *e.g. contacts are limited to those of the embedding social network, but contact type – casual, intimate, telephonic, etc. – may be attributes of the network of interest.*

Embedding fields and network display

- Embedding fields are the context in which the network exists.
- Some aspects of the context are relevant to the user's task, some are not.
- The display medium also can be considered as a hierarchy of embedding fields, the root of which is, say, the set of pixels of the display screen, intermediate levels might be 2-D and then 3-D spaces containing objects, while the leaves might consist of the coloured lines and objects used to show the network attributes of concern.
- The immediately ancestral embedding field for the display of the network may well be the appropriate environment in which to display the user-relevant contextual embedding field of the network.

Dynamic Properties of real networks

Network traffic changes over time, and networks themselves change.

If a network contains cycles, as most do, the traffic can vary regularly or chaotically.

The passage of traffic can alter the network stigmergically

- *e.g., in an infection network, the structure of the network changes when a node (person) moves from susceptible to infective to immune (or dead).*

Cycles are not possible in an infection network if persons become immune after being infected, even though the static structure of the network and its embedding field suggest that cycles should exist. Epidemic pulses must come from elsewhere – a larger network.

Framework: Categorizing Data Types

Six Descriptive Dimensions

from the Final Report of IST-013/RTG-002 (The HAT Report)

Acquisition	Streamed	Sporadic
		Regular
	Static	
Sources	Single	
	Multiple	
Choice	User-selected interactive	
	Externally imposed	
Identification	Located	
	Labelled	

Values	Analogue	scalar	
		vector	
	Categoric (crisp)	symbolic	linguistic
			Non-linguistic
		non-symbolic	linguistic
			Non-linguistic
	Categoric (fuzzy)	symbolic (non-linguistic)	
		non-symbolic (non-linguistic)	
Relations	User-structured		
	Source-structured		

Framework: Categorizing Display Techniques

One approach to categorizing: Survey and collate.

In parallel with the Framework Working Group, IST-059/RTG-025 has another Working Group developing an on-line Survey of network visualisation software. The survey is expected to be useful in its own right, but analysis of the properties of the surveyed items should also assist in developing the Framework categories.

Intuitive Categories

The Survey uses intuitively derived categories for describing the software. Some of those are obvious and irrelevant to the Framework, such as cost, open-source versus proprietary, hardware platform, coding language and extensibility, etc. Others are highly relevant, some being derived directly from the RM-Vis reference model.

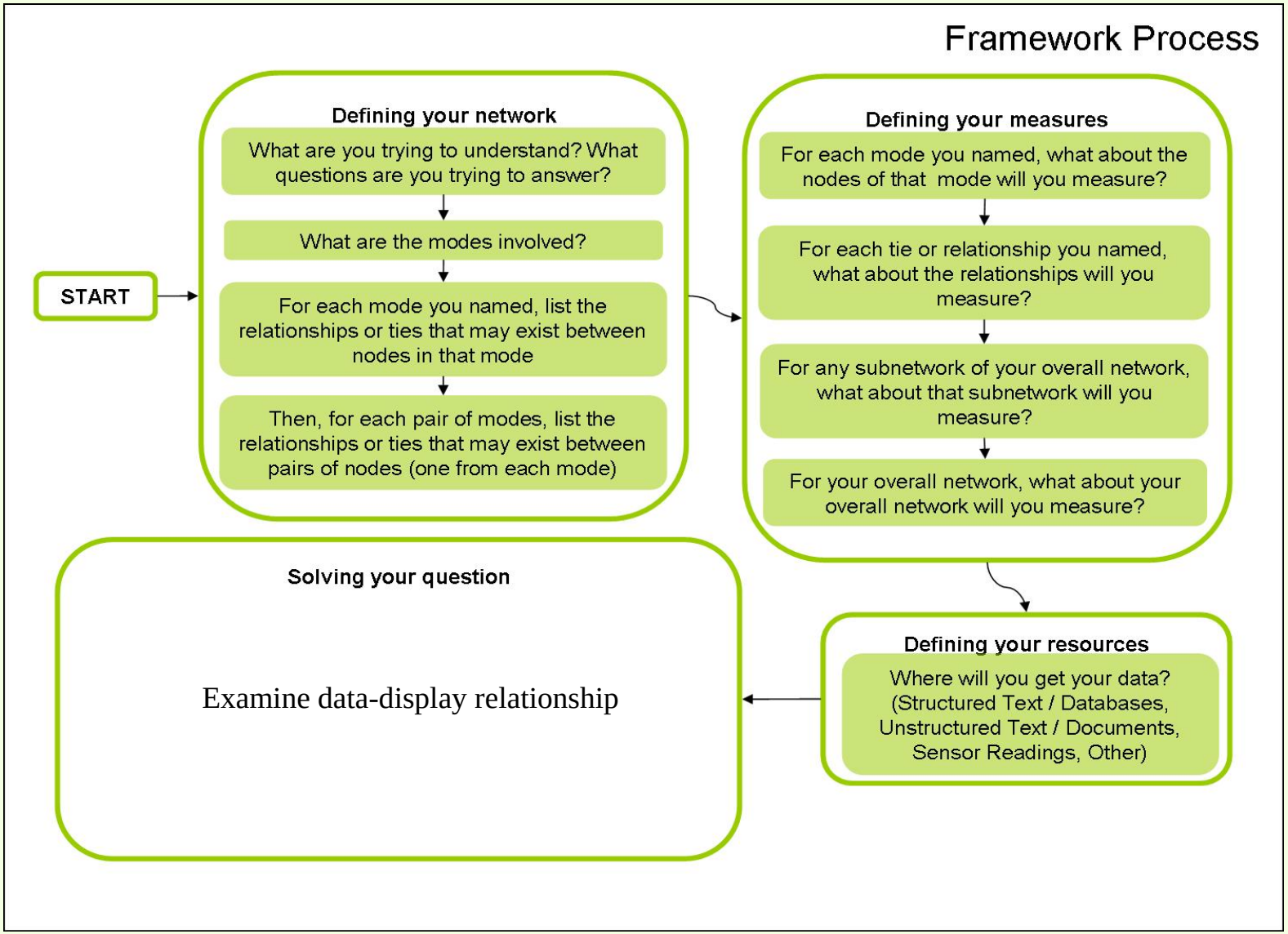
Framework: Categorizing Display Techniques

Four Descriptive Dimensions

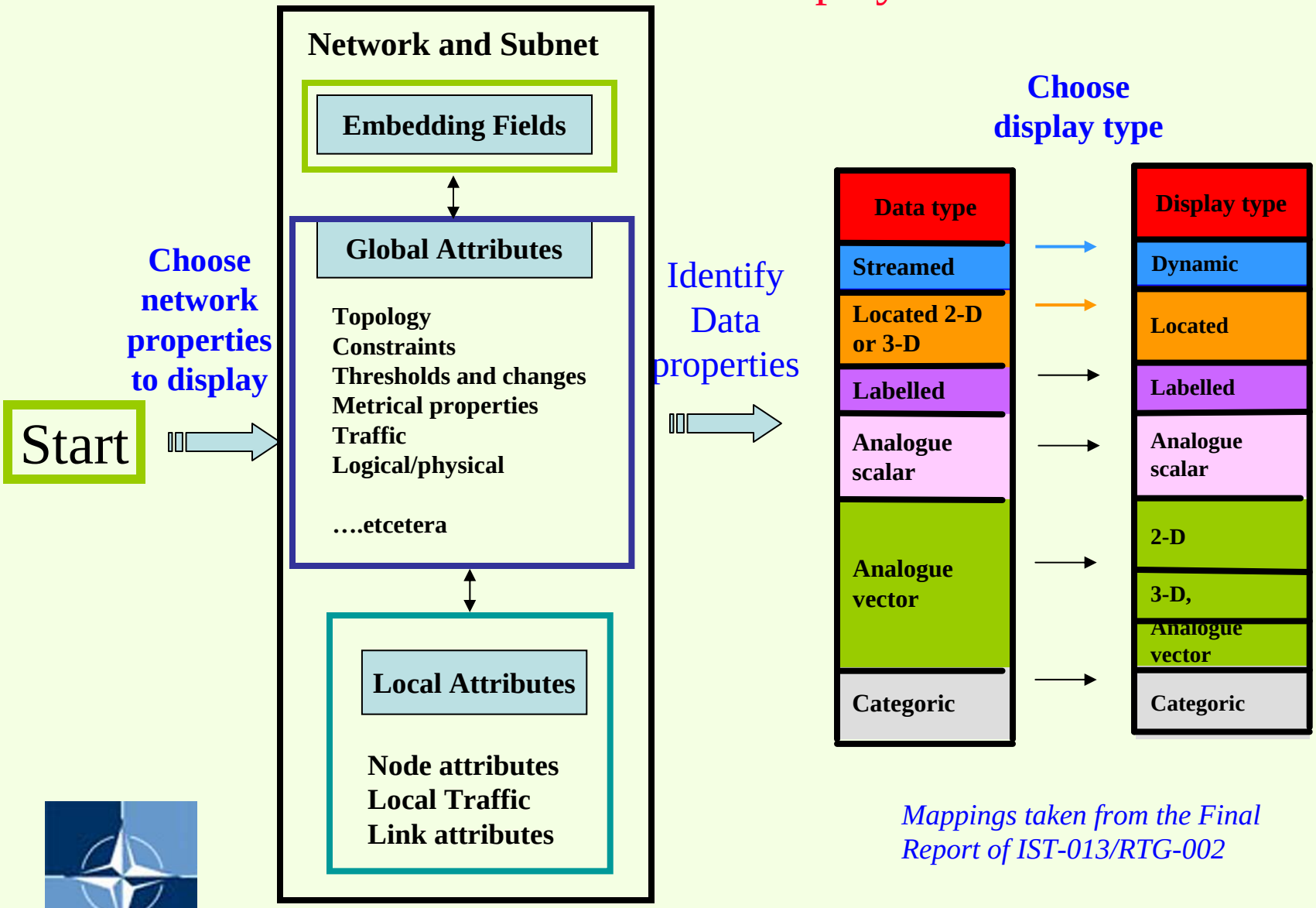
from the Final Report of IST-013/RTG-002 (The HAT Report)

Display Timing	Static picture	
	Dynamic variation	
Data Selection	User-selected interactive	
	Algorithmically selected	
Data Placement	Located	
	Labelled	
Data Values	Analogue	scalar
		vector
	Categoric	linguistic
		symbolic

Framework as Process: Network attributes



Framework as Process: Display attributes



Mappings taken from the Final Report of IST-013/RTG-002

Summary: Framework for Network Visualisation

•Many different kinds of network representation have been developed, but without a coherent foundation that would allow good representations to be used for other projects. **A good Framework provides that foundation**.

•A good **representation** supports the purposes of a **user** effectively.

•A Framework requires consideration of **both** the **user** and the **range of network properties** that might be **represented** in support of the **user's purposes**. Therefore a Framework must consider the nature of real networks as well as the properties of abstract mathematical graphs.

•Real networks are more **complicated** than are the abstract mathematical networks, though the mathematics remains relevant to the real networks.

•Real networks are often **fuzzy**. Links and nodes may be of variable quality. Nodes **transform** the kinds of traffic they receive and emit.

•Real networks are embedded in **user-relevant context** that affects their properties and behaviour. The context may itself be a network.

Framework — The Way Ahead

1. Complete the Framework by
   - Categorizing computable network attributes
   - Categorizing Network-related user tasks
   - Categorizing network-related display techniques
   - Develop mappings across categorizations:
     - task - attribute
     - attribute - display
2. Link the Framework with the Survey of Network Visualisation Software
3. Describe the Framework process for end users
   - Propose support software to guide the user in the Framework process
4. Test Framework use in different scenarios, and rework
5. Publish for general use.

IST-059/RTG-025 does not have the resources to complete all the above!

A Framework for Network Visualisation

Progress Report

Framework Working Group of IST-059/RTG-025

(J.-T. Bjørke, M. R. Nixon, M. M. Taylor, A. K.C.S. Vanderbilt, M. Varga)