ICCRTS 15

Paper 168

Assessing the Difficulty and Complexity of ELICIT Factoid Sets

May 2010

Anthony Alston

Centre for Applied System Studies,
Defence Academy of the UK,
Shrivenham, SW6 8LA, UK.

Tel. +44 (0)1793 585761

Email: a.alston@cranfield.ac.uk

1 Introduction

At the core of NEC [1] and NCW [2] is the tenet that greater information sharing results in greater shared awareness eventually leading to better decisions. The ultimate expression of this tenet is the Edge Organisation [3]; where information is shared to such an extent that those at the 'edge' of the organisation are empowered to act independently due to their high level of shared awareness. However, apart from some anecdotal evidence there is little empirical evidence to support this. To make good this shortfall the CCRP has commissioned an experimental platform called ELICIT[1] and has made it available free to researchers world wide, with the proviso that all experimental data is made available to the wider defence community through the CCRP [4].

ELICIT is currently being used to quantify the improvement attained through the use of different information sharing regimes, most notably in Portugal [5], which show that Edge Organisations do offer some advantages over traditional hierarchical information sharing structures. Dstl, the UK MOD science and technology agency, wishes to extend this work to include the solving of complex problems to test the hypothesis, 'The more complex the problem, the higher C2 maturity an organisation must be at to solve it.' However, currently there is no way of quantifying the difficulty of a problem (as presented in the form of a Factoid Set), either absolutely or relatively with one another.

Dstl has tasked the Cranfield University Centre for Applied Systems Studies (CfASS), based at the Defence Academy of the UK, to support the establishment of an experimental programme to investigate the hypothesis. The tasking on CfASS included:

1. Support the development of a Factoid set generation tool by investigating methods to quantify the complexity of a Factoid set.

2. The generation of a militarily realistic Factoid set representing a complex problem.

3. Suggest extensions to the ELICIT platform to make it more 'realistic' and enable it to handle different types of problem.

The body of this paper reports on the outcome of the first task. It initially discusses different problem-types and postulates a way of measuring their 'difficulty'. To test the usefulness of the method, the method is then applied to ELICIT and its current Factoid sets. The paper then finishes off with some concluding remarks. The second task is not presented here and the third task, is discussed in Annex A, particularly in the context of experimental set-up.

[1] ELICIT is a platform for the comparison of team problem-solving performance under different information-sharing regimes. At its heart are four Factoid Sets that provide the baseline data for the problem-solving exercises.

2 Assessing Problem Difficulty

2.1 Introduction

In assessing the difficulty of a problem it is necessary to consider two views of the problem being assessed; the view of the problem setter and the view of the problem solver. When developing a problem, the setter will consider the type of problem he is setting and also the experience level of the target solver. (This includes the context the setter expects the problem to be solved within. This is important and will be considered in greater detail later). For example, the setter for the Times crossword will be targeting a different solver community than, say, the setter of the crossword in a tabloid newspaper like the Sun. We will call this resultant difficulty the 'Target Problem Difficulty'.

The difficulty as perceived by the solver is dependent their own experience for that puzzle category and not upon the Target Problem Difficulty, For example, a solver used to the Times crossword would find the Sun crossword trivially easy even if the setter devised one that was difficult for someone from the Sun readership. We call this resultant difficulty the 'Perceived Problem Difficulty'.

The relationship between Target Problem Difficulty and Perceived Problem Difficulty is shown in Figure 1.

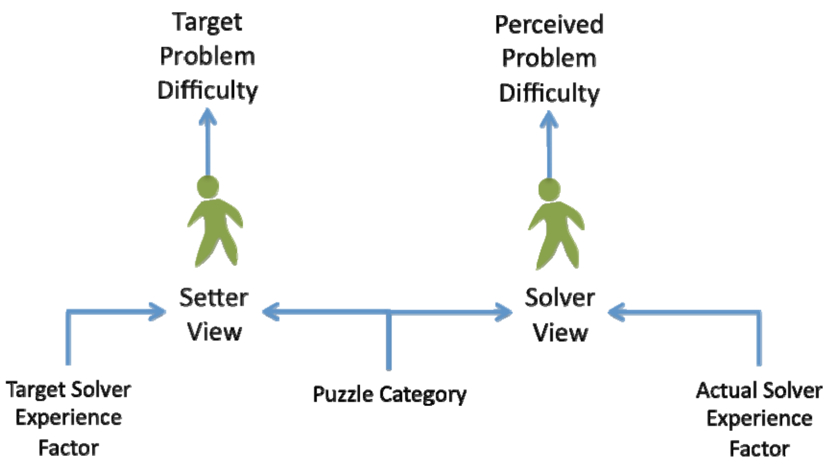

Figure 1: Problem Difficulty

An interesting point, shown within Figure 1, is that the Puzzle Category is the same from both the view of the Setter and the Solver. In the crossword example, the equivalent of the Puzzle Category is the crossword itself – it is the same puzzle regardless of who it is presented to for solution.

This sub-section will now discuss how to understand Puzzle Difficulty.

2.2 Puzzle Difficulty

Difficulty can be described in terms of three dimensions, which relate back to the two constituents (Puzzle Category and Experience) in the model of Figure 1:

- Problem Category. Problems can be categorised in many ways, with many different factors. To keep problem categorisation simple this paper groups the factors into two dimensions; ***Problem Type***, describing the nature of the problem, and ***Problem Complexity***, describing how the data/information relating to the problem/solution are represented. These dimensions are shown in Figure 2.

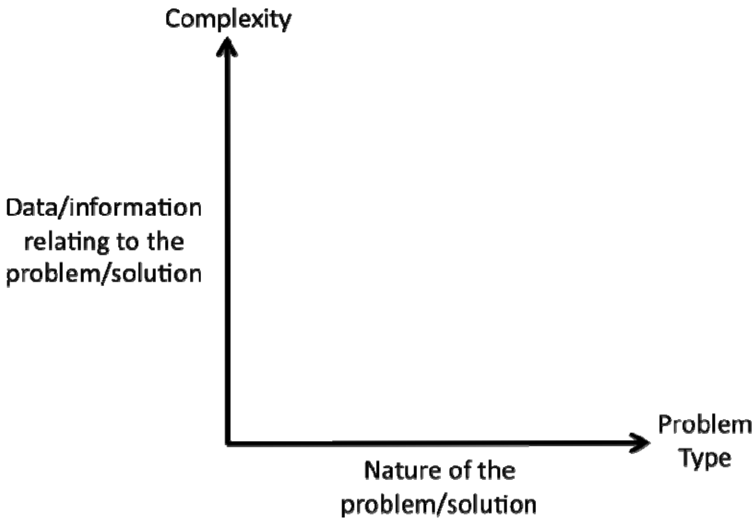

Figure 2: Derivation of Problem Category

- ***Experience***. Experience is also a combination of factors including the expertise of a particular solver for that type of puzzle and the environment the puzzle is solved within. As an example, school exams are more difficult if performed in a noisy room rather than a quiet one and are easier if the use of a calculator is allowed. In both these examples, the exam itself has not changed, but its perceived difficulty has.

The remainder of this section will now discuss these three dimensions: Problem Type, Problem Complexity and Experience; and then discusses the use of the resulting Problem Difficulty cube.

2.2.1 Problem Type

The definitions for the extreme values of Problem Type used within this paper are those devised by Rittel and Webber in 1973, who divided problems into 'Tame' and 'Wicked' [6]. They distinguished between them thus:

> "The kinds of problems that planners deal with--societal problems--are inherently different from the problems that scientists and perhaps some classes of engineers deal with. Planning problems are inherently wicked.
>
> As distinguished from problems in the natural sciences, which are definable and separable and may have solutions that are findable, the problems of governmental planning--and especially those of social or policy planning--are ill-defined; and they rely upon elusive political judgment for resolution. (Not "solution." Social problems are never solved. At best they are only re-solved--over and over again.)

> The problems that scientists and engineers have usually focused upon are mostly "tame" or "benign" ones. As an example, consider a problem of mathematics, such as solving an equation; or the task of an organic chemist in analyzing the structure of some unknown compound; or that of the chess-player attempting to accomplish checkmate in five moves. For each the mission is clear. It is clear, in turn, whether or not the problems have been solved.
>
> Wicked problems, in contrast, have neither of these clarifying traits; and they include nearly all public policy issues--whether the question concerns the location of a freeway, the adjustment of a tax rate, the modification of school curricula, or the confrontation of crime."

For a Tame problem, the main distinguishing features, as described above, are "For each [Tame problem] the mission is clear. It is clear, in turn, whether or not the problems have been solved."

The concept of Wicked Problems was defined as being those that:

- Describe social planning problems that are difficult or impossible to solve because of incomplete, contradictory, and changing requirements
- Are often difficult to recognize and because of complex interdependencies, the effort to solve one aspect of a wicked problem may reveal or create other problems

Wicked Problems have ten characteristics. The two that are most relevant to this work are:

> 'There is no immediate and no ultimate test of a solution to a wicked problem.'

and

> 'Every wicked problem can be considered to be a symptom of another problem.'

This suggests that a major defining aspect of Wicked Problem is intricate way that the problem and its solution are inter-twined, with the one solution leading onto another problem. This means that it is not possible to determine whether the problem will ever end, and hence whether there is or is not a solution at all. These values are shown on Figure 3.

2.2.2 Problem Complexity

The second dimension of Puzzle Categorisation relates to the representation of the data/information relating to the problem. Here the concept of Complexity is used to grade this dimension.

One of the principal characteristics of complexity is the number and nature of the inter-relationships between the elements of the system or problem under investigation [7]. At one extreme of the complexity dimension the number of elements is small and their inter-relationships few and well understood. Within this paper this state is referred to as 'simple'. At the other extreme of the dimension a problem may have many elements, with a large and dynamic set of non-deterministic inter-relationships. In this paper this state is referred to as 'Complex'. These values are shown in Figure 3.

2.2.3 Area of Puzzle Feasibility

When the dimensions of Problem Type and Problem Complexity are brought together it becomes obvious that not all points on the surface composed are viable. This is shown in Figure 3.

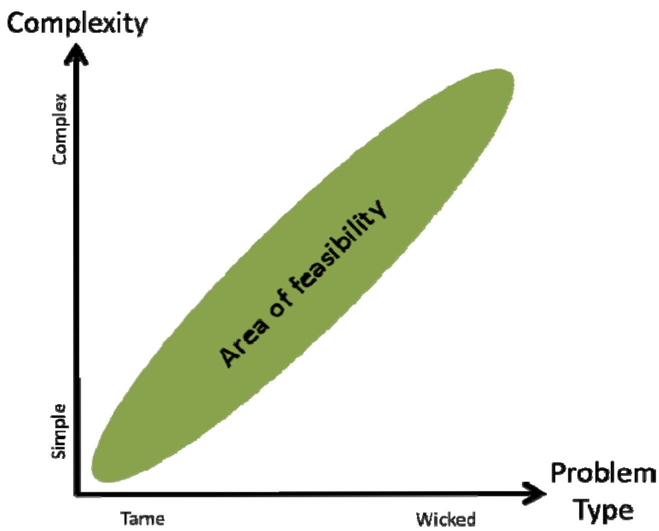

Figure 3: Problem Categorisation – Area of Feasibility

The Figure shows that there is an Area of Feasibility where viable problems reside. The rationale for the shape of this area is:

- If a Tame problem is one in which there is a single solution that is obvious when it has been reached, then it cannot have elements whose inter-relationships are dynamic and non-deterministic.
- If a problem is simple, it has few elements whose inter-relationships are few and well understood, then it cannot also be wicked.

2.2.4 Experience

The third dimension that defines Problem Difficulty is that of Experience. The Experience Dimension, like the other dimensions, has many factors but all relate to the suitability of the solver to solve the problem and includes the structure of the solver (i.e. is one person, a team, the team structure, whether the team has worked together before) as well as their previous knowledge of that type and complexity of problem. This provides a three-dimensional feasibility space as shown in Figure 4.

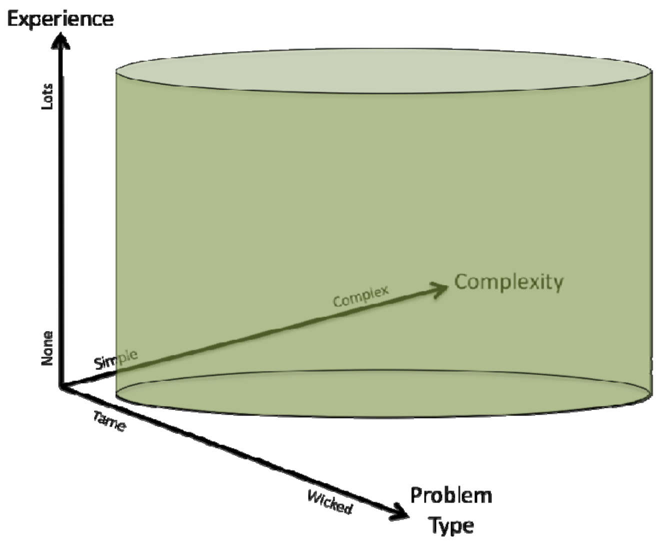

Figure 4: Problem Category – Feasible Volume

2.2.4 The Problem Difficulty Cube

It can be assumed that the lesser the experience the more difficult the problem will seem. This is shown in Figure 5.

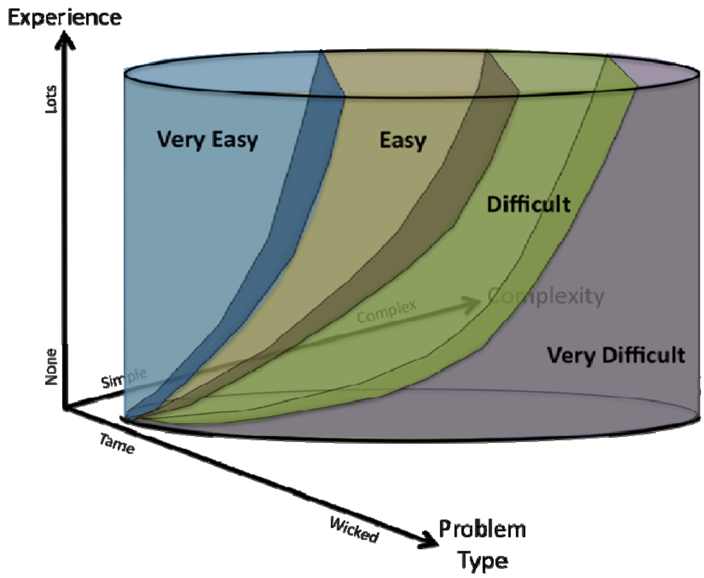

Figure 5: The Problem Difficulty Volume

The figure shows distinct bands of difficulty (This is for illustrative purposes only to show the profile of difficulty within the space). The shape of the bands shows how, for any level of experience, the Puzzle Difficulty rises as the problem becomes more Wicked and Complex.

This space can be used to indicate the difference between Target Problem Difficulty and Perceived Problem Difficulty that were illustrated in Figure 1, and this is shown in Figure 6. In this Figure, the constant problem category is shown as the line resulting from the meeting of the problem's complexity and type. A point on the vertical Problem Category line indicates the difficulty of the problem where it intercepts a line from the Experience axis. In the figure, two experience levels are shown, Experience B representing a solver of low experience and Experience A a solver of higher experience. The difficulty rating shown rises from 'Easy' to on the border of 'Difficult' and Very Difficult' for the two experience levels.

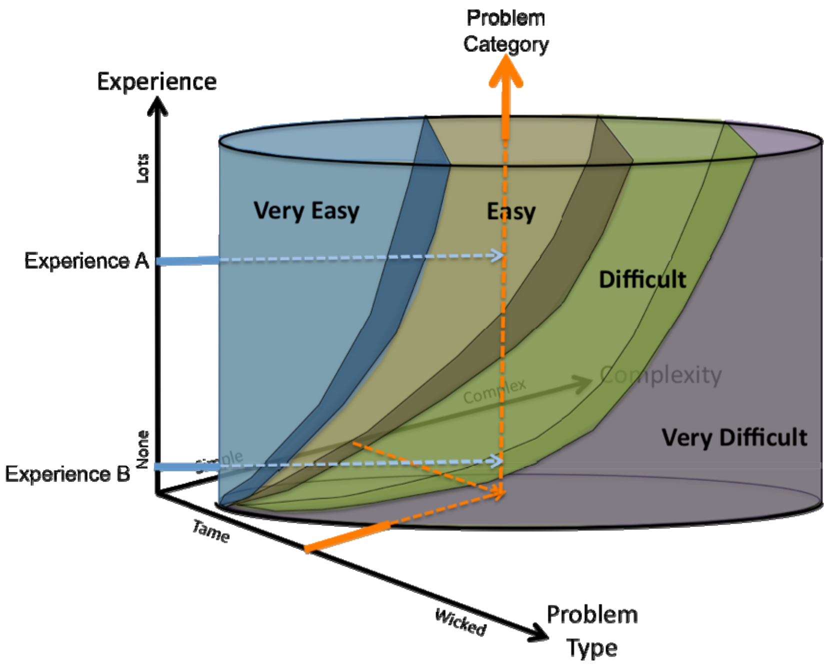

Figure 6: Difference in Difficulty of different Experience Factors

The actual values on the axes are not absolute; they are relative to the context of the problem. The next two sections of this paper take this Problem Difficulty space and firstly identifies the part of it that is relevant to problems within the ELICIT platform and identifies the parameters for the axes when applied to ELICIT, then lastly a brief analysis of the four, standard ELICIT Factoid sets[2] are shown.

3 Measuring Problem Difficulty in ELICIT and its Factoid Sets

The ELICIT platform has been developed to investigate very specific issues: how does changing the information sharing regimes affect the 'shared awareness' of the problem-solving team and do these different information sharing regimes impact the problem-solving performance of the team [10].

Nominally, ELICIT accommodates a team of seventeen players who are presented with a problem which they have to solve. Performance is measured in terms of the number of players who deduce the correct solution (a proxy for a measure of shared awareness) and the time taken to deduce the solution (a proxy for the effectiveness of the information sharing regime against that problem). The problem is presented to the players in terms of a Factoid Set.

A Factoid Set is a collection of statements which, when logically combined, allows the players to deduce the solution to the problem. A Factoid set contains nominally sixty-four Factoids that are either 'key' to solving the problem (the minimal sub-set that can be used to solve the problem), 'supporting' (they agree with the solution but on their own cannot be used to deduce the solution) and

---

[2] ELICIT, as provided by the CCRP, has associated with it four factoid sets as standard for use by the experimental community.

'noise' (irrelevant to, but not contradictory with, the solution). ELICIT allows the problem setter to specify who receives which Factoid and when a Factoid is distributed.

Factoids can be shared between the players either by a player deliberately 'pushing' a Factoid to another player or 'publishing' the Factoid to one of four web-sites, access to which is controlled by the problem setter.

The answer to the problem is of the form of a 'who', 'what', 'where' and 'when' of a terrorist incident. Any player can make a guess at any time.

In order to measure the performance of the team, in the way indicated above, the problem represented by the Factoid Set has just one known solution and the statements within the Factoid Set are consistent with each other and complete – all the data required to deduce the solution are present. This means that ELICIT can only be used to investigate deductive problems where the solution is known.

The above description of the ELICIT platform and its associated Factoid Set provide the following restrictions on Problem Characterisation:

- Problem Type: The problem within ELICIT must have a solution, and the answer can be shown to be the correct one. Hence ELICIT as currently conceived can only deal with Tame problem types, and hence there is no opportunity to vary this variable.
- Complexity: The Factoid set contains a known number of Factoids, that are complete, not contradictory and whose logic relationships are such that the solution can be deduced unambiguously. This means that ELICIT can only manage puzzles towards the Simple end of complexity scale. However, there is an opportunity to set the complexity of a Factoid Set within this lower end of the scale through varying the number of Factoids required to deduce a solution and their logical relationships.

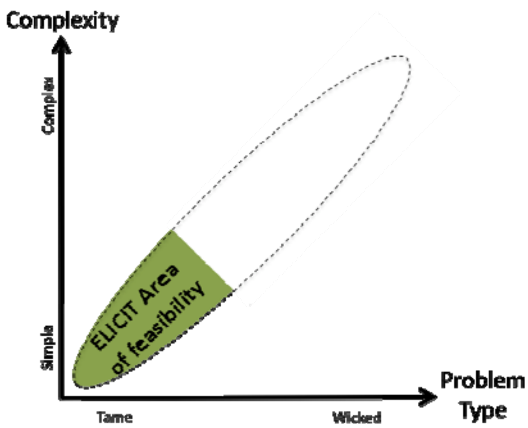

Figure 7: Problem Category

Figure 7 summaries this.

However, puzzle difficulty is related to the experience of the solver. To derive this requires analysis of the solver and the environment in which it is being solved. This analysis is beyond the scope of this paper, although issues relating to 'experience' are discussed in Annex A.

4 Measuring the Complexity of a Factoid Set

4.1 Variables for measuring the Complexity

The Problem Category for ELICIT is restricted to the area shown in Figure 7. Of the two dimensions, Problem Type is confined to Tame problems and cannot be varied. That leaves the only variables that can be changed to affect the Problem Category are those relating to Complexity. For an ELICIT Factoid set these variables are summarised in Table 1 below, and explained in more detail in the subsequent sub-sections.

Variable Category	Description	Variables
Conceptual Level	The 'structure' of the puzzle; how the logic for the sub-solutions interact with each other.	• Number of separate 'logic

		'logic chains' . • Interaction between the 'logic chains'.
Logic Level	The logical structure that has been employed to form a sub-solution.	• Number of Factoids required to deduce a sub-solution. • Number of relationships between the Factoids. • Number of candidate sub-solutions.
Factoid Level	The language used within a Factoid (for example clarity/vagueness of language used and the ease with which it is possible to identify which sub-solution it contributes to.	• Positive or negative language. • Clarity of Factoid. • Single or multiple Factoid Categorisation. • Factoid Categorisation not deducible from the Factoid language. • Sub-solution within Factoid. • Number of Noise Factoids.

Table 1: 'Complexity' Variables

4.1.1 Conceptual Level

The Conceptual Level variables measure the complexity of the Factoid set as a whole. It assumes that the solution has a number of sub-solutions, in the case of the current ELICIT Factoids there are four sub-solutions; 'who', 'what', 'where' and 'when'. The complexity at the Conceptual Level rises when the interaction between the logic arguments for each sub-solution rises. The Conceptual Level has two variables.

- **Variable: Number of Separate 'logic chains'**. The variable pertains to the form of the logical argument to deduce a sub-solution. The simplest measure is when there is a single logical argument for each sub-solution. The more logical arguments for a sub-solution then the more complex the problem.
- **Variable: Interaction between the 'logic chains'**. This variable measures the amount of interaction between the logical arguments ranging from the simplest, where there is no interaction, to the most complex where two or more arguments are intricately cross-linked. The cross-linking is in terms of common Factoids between the arguments.

4.1.2 Logical Level

The Logical Level variables measure the complexity of a logical argument. This Logical Level contains three variables:

- **Variable: Number of Factoids required to deduce a sub-solution**. It assumes that the number of Factoids within a logical argument is a proxy for the complexity of the sub-solution argument; the more the Factoids the more the complex the problem.
- **Variable: Number of relationships between the Factoids.** The number of relationships between the Factoids required to deduce a sub-solution is a proxy for the complexity of the logical argument; the more relationships the more complex the problem.
- **Variable: Number of candidate answers for a sub-solutions.** This variable assumes that the more candidate answers for a sub-solution there are the more complex the problem. For example, in the problem of who lives in which house a problem with six potential residents is more complex than one with four potential residents.

4.1.3 Factoid Level

The Factoid Level variables measure the complexity within a Factoid. The Factoid Level contains six variables:

- **Variable: Positive or Negative Language.** This variable measures the 'tense' of the Factoid. It assumes that a Factoid using positive language, 'Abi lives next door to Bertie', is less complex than a Factoid using negative language, 'Abi does not live in the red house'.

- **Variable: Clarity of Factoid**. This variable measures the clarity of the language used within the Factoid and relates to the certainty with which the solver can rely on the information contained to help solve the problem. For example, 'Abi lives next door to Bertie' is clearer, and hence simpler than, 'Abi might live next door to Bertie', which is vague and hence is more complex.

- **Variable: Single or multiple Categorisation**. This variable measures the number of sub-solutions, categories, that are contained within a Factoid. It assumes that the more it contains the more complex the problem. For example, 'Abi lives in the street', relates only to the 'who' sub-solution, whilst 'Abi lives next door to the red house' relates to the 'who' and 'where' sub-solutions.

- **Variable: Factoid Categorisation not deducible**. This variable assumes that if a Factoid contains information pertaining to a category (e.g. 'who') then it is needed to solve the respective sub-solution. However, it is assumed that the solution is more complex to deduce if a Factoid required to solve a sub-solution does not contain information of that categorisation.

- **Variable: Sub-solution within a Factoid**. Most Factoids contain information that must be combined logically with other Factoids to deduce a sub-solution. However, the setter can provide the solver with short-cuts by providing sub-solutions or partial sub-solutions within a Factoid. It is assumed that doing this simplifies the problem.

- **Variable: Number of noise Factoids**. A Noise Factoid has already been defined as a Factoid containing irrelevant, but not contradictory information. It is assumed by this measure that the more Noise Factoids that a solver has to filter-out the more complex the problem.

4.2 Measuring the difficulty of the ELICIT Factoids Sets

The above metrics and variables were tested against the four, standard Factoid sets provided by the CCRP with the ELICIT platform. The following is just the results of this analysis. The author has decided that providing the details of the analysis in this open paper could undermine the use of the Factoid sets within experiments; effectively it would be a 'spoiler' providing the answers. The author would be happy to provide this analysis on request.

Some of the variables identified above are of little use in assessing the variation of difficulty across the four ELICIT Factoid Sets as they do not explicitly change. Table 2 below shows those that have been used in the assessment.

Variables	Comments

Conceptual Level 1.Mixed logic streams 2.Number of logic streams per sub-solution	All the four Factoid sets have single logical streams per sub-solution. Hence the 'number of logical streams' variable is of no use in the assessment.
Logic Level 1.Number of Factoids required to deduce sub-solution 2.Number of relationships between the Factoids 3.Number of candidate sub-solution 4.Number of Factoids	The number of candidate sub-solutions is the same for each Factoid Set.
Factoid level 1.Factoid Categoriation not deducible from the Factoid language 2.Compound Factoid (or) 3.Clarity of Factoid - Defines key or support Factoid 4.Positive or negative language 5.Compound Factoid (and) 6.Single or multiple Factoid Categorisation 7.Noise Factoids 8.Sub-solution within Factoid	The structure of the Factoids is re-used across the Sets. All these variables are therefore constant across the Factoids Sets.

Table 2: Variables used to assess the complexity of the Factoid Sets

complexity of the Factoid Sets

Table 3 below shows the results of analysing the four Factoid Sets using the variables identified within Table 2. The full analysis is not presented here, but can be made available on request to the author.

	Factoid Set 1	Factoid Set 2	Factoid Set 3	Factoid Set 4
Mixed Logic Streams:	7	8	9	4
Factoids per sub-solution:	5,5,5,9 (24)	5,11,8,10 (34)	10,8,14,4 (36)	5,7,6,4 (22)
Number of relationships:	25	25	27	17
Number of Factoids:	15	15	16	12

Table 3: Results of the Analysis of the Factoids Sets

The analysis was done on the minimum sub-set of Factoids required to deduce the four sub-solutions. Looking at each variable in turn:

- **Mixed Logic Streams**: This is measured by counting the number of logical relationships that cross between the logic streams for each sub-solution. The values for Factoid Sets 1 to 3 are approximately the same, with Factoid Set 4 being less.

- **Factoids per sub-solution**: This variable is measured by counting the number of factoids required to solve each sub-solution (Who, What, Where , When) and then the total these. When a sub-solution is required as part of another logic stream (for example, if the Who sub-solution is required before the What sub-solution can be deduced) then the number of preceding Factoids are also included in the count. The analysis shows that Factoid Sets 2 and 3 are comparable, with Factoid Sets 1 and 4 also comparable but with considerably less Factoids required to obtain the solutions.

- **Number of Relationships**: This variable is measured by counting all the relationships between the Factoids. The analysis shows comparable values for Factoids 1 to 3, but considerably fewer for Factoids Set 4.

- **Number of Factoids**: This variable is the number of Factoids in the sub-set required to find all the sub-solutions. The analysis shows comparable values for Factoid Sets 1 to 3, with fewer for Factoid Set 4.

Table 4 and 5 are an attempt to quantify the complexity of the Factoid Sets. Table 4 assigns the ordinal value to each Factoid Set for each variable, and then adds them up. Table 5 simply adds up the figures shown in Table 3.

	Factoid Set 1	Factoid Set 2	Factoid Set 3	Factoid Set 4
Mixed Logic Streams:	3	2	1	4
Factoids per sub-solution:	3	2	1	4
Number of relationships:	2=	2=	1	4
Number of Factoids:	2=	2=	1	4
TOTAL	***10***	***8***	***4***	***16***

Table 4: Ordering by order within each variable

	Factoid Set 1	Factoid Set 2	Factoid Set 3	Factoid Set 4
Mixed Logic Streams:	7	8	9	4
Factoids per sub-solution:	24	34	36	22
Number of relationships:	25	25	27	17
Number of Factoids:	15	15	16	12
TOTAL	***71***	***82***	***88***	***55***

Table 5: Ordering by totalling value of each variable

Unsurprisingly the two tables provide consistent results. Factoid Set 3 is clearly the most complex, with Table 5 showing Factoid Set 2 nearer in complexity to Factoid Set 3 than Table 4. Both show Factoid Set 4 the least complex by far.

5 Summary and conclusions

ELICIT is an experimental platform for the investigation of the impact of different information sharing regimes on team-problem solving. However, currently there is no means of measuring the difficulty, or complexity, of the Factoid sets that constitute the data sets that present the problem to the teams under investigation. This paper presents a general model for measuring difficulty of different problem types, and from this a method of measuring the difficulty and complexity of an ELICIT Factoid Set.

The general method of measuring difficulty has three variables:

1. Problem Type: whether the problem is tame (has a single, agreed answer) or is wicked (has no right answer, and any proposed solution may have unknown, knock-on consequences).

2. Problem complexity: the number and nature of the facts required to provide a solution and the number and nature of the relationships between the facts. At one end of the scale the facts and their relationships are few and non-contradictory. At the other end of the scale there are many contradictory facts with difficult to understand relationships.

3. Experience: A measure of the 'solver's' previous knowledge of this type of problem and the environment in which the problem is being solved.

From examining these variables it is obvious that without knowing the experience of the teams the Factoid sets can only be measured for their intrinsic difficulty established by the problem setter; and this has been undertaken for the four Factoid sets distributed by the CCRP with the ELICIT platform (however, a discussion about experience and experimental set-up is provided in Annex A).

In conclusion, the method as used does enable the distinction between the different four Factoid sets, but how well will only be determined when the measurements are correlated with empirical data derived from actual experimental runs. The actual analysis data has not been presented here as this contains the solution to the four Factoid sets, potentially rendering them unusable in an experimental context. However, the author is happy to distribute the results on request.

A next step following the generation of the measurement method would be to see if it could be used as part of generative process, hence ensure a problem of given difficulty is produced.

6 References

[1] NEC, Understanding Network Enabled Capability, 2009, Newsdesk Communications

[2] Network Centric Warfare, Alberts et al, 2000, CCRP.

[3] Power to the Edge, Alberts and Hayes, 2003, CCRP

[4] ELICIT – The Experimental Laboratory for Investigating Collaboration, Information-sharing and Trust, Rudd, 2007, ICCRTS 12.

[5] ELICIT and the Future C2: Theoretical Foundations for the Analysis of ELICIT Experiments, Manso and Nunes, 2008, ICCRTS 13.

[6] Dilemmas in a General Theory of Planning, Rittel and Webber, 1973, Policy Sciences (4).

[7] Complexity Theory and Network Centric Warfare, Moffat, 2003, CCRP

Annex A: Extending the ELICIT Platform

A1 Introduction

The Dstl tasked The Centre for Applied Systems Studies (CfASS) at Cranfield University to produce a Complex Factoid Set within a realtistic military context for the ELICIT experimental platform. As was discussed in the main body of the report, it is not possible to generate a truly complex Factoid Set due to three reasons:

1. The ELICIT Platform restricts the experiments that can be performed.
2. The Factoid Sets are for a problem that is tame.
3. The metrics used restrict the experimentation to investigating tame problems.

This Annex discusses potential extensions to the ELICIT platform by considering how the experimental set-up could be extended to enable the investigation of matching different forms of information sharing (and C2 approaches) to the complexity of the problem.

A2 The ELICIT experimental set-up

The experimental set-up can be divided into three sections:

1. The Structure of the Team solving the Problem. This includes the roles and responsibilities of the groupings within the team, their 'command' relationships and the information sharing regimes.
2. The Structure of the Problem as presented. In our discussion this is the Factoid set.
3. The structure of the solution. This is the form of the solution required by the setter.

Currently these three areas are aligned as shown in Figure A1.

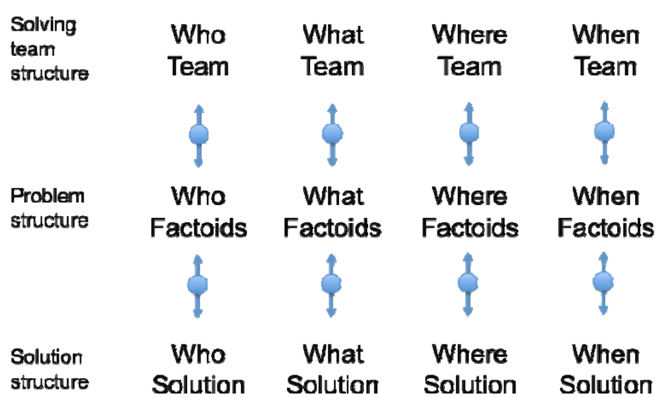

Figure A1: The current experimental set-up

This alignment means that the groups know which part of the problem they are responsible for, and which parts others are responsible for. This means:

1. Groups know which Factoids to keep and which to pass on, and to whom.
2. They know when their part of the problem has been solved.

The tasking on CfASS was to make a more realistic military, complex Factoid set within this structure. The simplest step towards producing such a Factoid Set was to re-write the Factoids in 'military speak' within a military context, which is not presented here.

The next step after producing a more realistic Factoid set was to change the Team structure to one that better represents a military organisation. The result of this, shown in Figure A2.

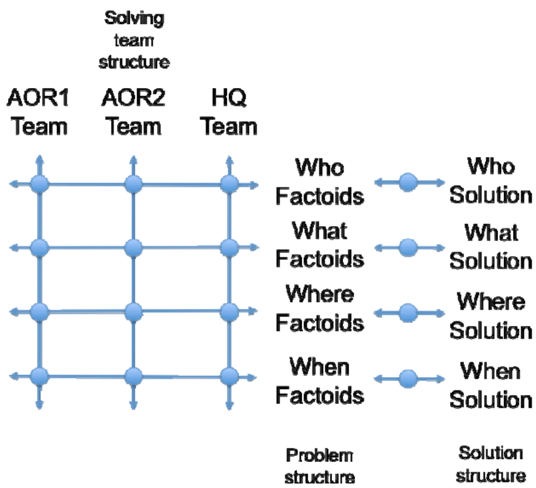

Figure A2: Single Twist Structure

Now it is no longer possible for the groups within the team to know which part of the problem they are responsible for, as the team structure no longer mimics the problem structure. This can be thought as putting a 'single twist' in the experimental set-up. Putting this single-twist in has a number of consequences:

1. Groups have no longer know which part of the problem they are responsible for, so do not know which Factoids are relevant and which they should pass on.
2. The Groups AOR1 and AOR2 could use the Factoids available to them and speculate about the solution and pass these speculations to the HQ. Here, perhaps with the aid of further Factoids, a solution could be 'guessed'. This arrangement makes the problem far more complex, though perhaps no longer a deducible problem. However, this arrangement also requires:
    a. 'Speculations' to be passed between participants.
    b. Does not require Factoids to be passed.

This is a great departure from the original experimental set-up. There is one more step that can be taken and this is represented in Figure A3.

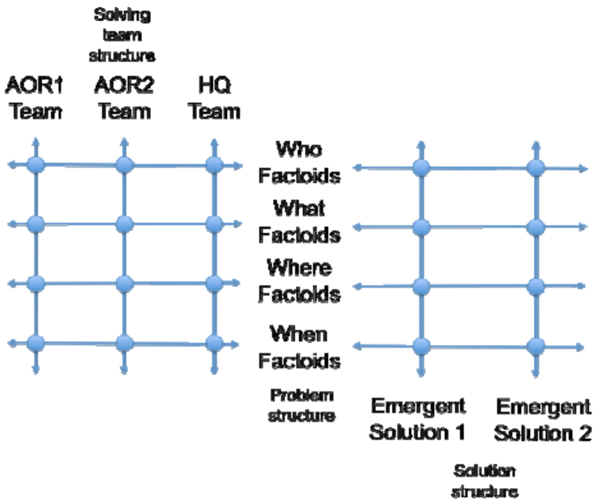

Figure A3: Double Twist Structure

Figure A3 shows all three elements of the experimental set-up mis-aligned. Now the structure of the Factoid Set has no relevance to the solution structure or the Team Structure. An example of such an experimental set-up would be a desired solution of 'achieving stability in town X' with the Factoids relating to 'food and water supply', 'insurgent activity' and 'cost of maintaining operations'. This is now a truly complex/wicked problem, with no single 'right solution' and a non-determinate Factoid structure. This type of problem requires:

1. The capability for the groups and individuals to 'chat' between each other.
2. The HQ to determine when the 'solution' has been attained. (That is, it is no longer a matter of the HQ guessing the answer, but the HQ saying 'this is the best we can do'.

A3 Conclusion

There are some 'cosmetic' extensions to the ELICIT platform that would make it more realistic, for example the addition of a graphics front end and the use of a more realistic military Factoid Set. The risk is that making the 'ELICIT experience' more realistic only opens it up to unfavourable comparison with 'the real world'; something which could not be level against the current, very stylised and somewhat artificial, Factoid Set.

This Annex has tried to explore how to make the problem more complex/wicked, by changing the experimental set-up. This would require:

1. The ability to pass 'assumptions' between the players. This could be done relatively easily within the current ELICIT platform by extending the use of the existing 'guess' facility to allow the guess to be sent to other players in exactly the way Factoids are passed.
2. More complicated metrication and measuring, and information sharing regimes.

The allure of the double-twist experimental set-up is that it not only provides the potential for a very realistic scenario, but also the ability to much more accurately mimic many different types of C2 approaches and problem-solving structures and distinguishing between subtle differences in behaviour. This is due to the ability to accurately reproduce structures for problem-solving responsibilities as well as different information sharing regimes. However, the main cost of this lies in an increase in the effort required to set-up, design and run the experiments.