DEFENCE **R&D** DÉFENSE

Development of a Team Scenario Content Generation Framework

Andrew Morton

Barbara D. Adams

Defence R&D Canada
Contract Report
DRDC Toronto CR 2010-105
March 2010

Canada

DEVELOPMENT OF A TEAM SCENARIO CONTENT GENERATION FRAMEWORK

by:

Andrew Morton and Barbara D. Adams

Human*systems*® Incorporated
111 Farquhar Street
Guelph, ON N1H 3N4

Project Manager:

Barbara D. Adams
(519) 836-5911

PWGSC Contract No. W7711-088128/001/TOR
8128-11

On Behalf of

Department of National Defence

As represented by

Defence Research and Development Canada Toronto
Toronto, Ontario, Canada
M3M 3B9

DRDC Toronto Scientific Authority:

Dr. Matthew Duncan
(416) 635-2000 (ext. 2020)

March 2010

Author

Original signed by Barbara D. Adams

Barbara D. Adams
Human*systems*® Incorporated

Approved by

Original signed by Dr. Matthew Duncan

Dr. Matthew Duncan
Head, Collaborative Performance and Learning Section

Approved for release by

Original signed by Dr. Joseph V. Baranski

Dr. Joseph V. Baranski
Chair, Knowledge and Information Management Committee
Chief Scientist

,

Abstract

Web-based software called ELICIT (Experimental Laboratory for the Investigation of Collaboration, Information-sharing, and Trust; Parity Communications Inc., 2008) offers a research platform for exploring team performance dynamics, and command and control (C2) organizations. This project aimed to use existing ELICIT scenarios as a starting point for developing a formal relational/structural framework that would enable the creation of a means of systematically generating the content required for complex team scenarios. This knowledge could be instructive for designing a workable approach to use for tightly controlled future scenario generation.

ELICIT scenarios contain 68 short statements, typically a single sentence, called factoids. Analyses conducted in this course of this project slowly progressed from decomposition of the ELICIT factoids into bits of information, to the level of factoids as a whole, to the level of the scenario as a whole. A number of key observations were made through this process.

The primary observation is that while the scenarios have an underlying structure which allows both deductive reasoning and inference, there seems to be little ability for the systematic creation of factoids or manipulation of scenario difficulty. Through the analyses conducted there appears to be a basic scenario structure upon which a framework for the understanding of difficulty, complexity, and other potentially interesting experimental manipulations could be built. The analyses of this project represent a start toward the goal of better understanding (and hence controlling) scenario characteristics and systematic scenario creation.

Résume

Le logiciel web portant le nom d'ELICIT (*Experimental Laboratory for the Investigation of Collaboration, Information-sharing, and Trust*, de Parity Communications Inc., 2008) offre une plateforme de recherche pour l'exploration de la dynamique de la performance collective et des organisations de commandement et contrôle (C2). Ce projet visait à utiliser les scénarios ELICIT existants comme point de départ pour l'élaboration d'un cadre relationnel/structurel formel qui permettrait la création d'un moyen de créer systématiquement le contenu nécessaire pour les scénarios d'équipe de type complexe. Cette connaissance pourrait être instructive pour concevoir une approche apte à être mise en pratique pouvant être utilisée pour la production future de scénarios strictement contrôlés.

Les scénarios ELICIT contiennent 68 courts énoncés, habituellement formulés en une seule phrase, appelés « factoïdes ». Des analyses menées au cours de ce projet ont progressées lentement, de la décomposition des factoïdes d'ELICIT en segments d'information jusqu'au niveau des factoïdes dans leur ensemble, puis au niveau du scénario dans son ensemble. Un certain nombre d'observations clés ont été faites au cours de ce processus.

La principale observation touche le fait que bien que les scénarios possèdent une structure sous-jacente qui permet de faire à la fois des raisonnements déductifs et de l'inférence, la création systématique de factoïdes et la manipulation du niveau de difficulté des scénarios semblent difficiles. Les analyses effectuées ont permis de dégager ce qui semble être une structure de scénario de base à partir de laquelle il serait possible de bâtir un cadre permettant de comprendre la difficulté et la complexité des scénarios ainsi que d'effectuer d'autres manipulations expérimentales potentiellement intéressantes. Les analyses de ce projet constituent un point de départ vers une meilleure compréhension (et ainsi un meilleur contrôle) des caractéristiques d'un scénario et de la création systématique de scénarios.

Executive Summary

Development of a Team Scenario Content Generation Framework

Andrew Morton and Barbara D. Adams, Human*systems*® Incorporated; DRDC Toronto No. CR 2010-105; Defence R&D Canada – Toronto; March 2010

Studying team performance dynamics in the laboratory requires an experimental platform. One such platform is ELICIT, the Experimental Laboratory for the Investigation of Collaboration, Information-sharing, and Trust (Parity Communications Inc., 2008). ELICIT was developed to experimentally study command and control (C2) organization, team and individual characteristics, and the network-centric value chain. ELICIT is a web-based software that is predicated on the general scenario that a terrorist attack will occur at some point in the future within a hypothetical situation, and teams of up to 17 members must work together to determine the details (i.e. Who, What, Where, and When) of the pending attack. To do this, each member of the "team" receives distinct units of information, and they must choose what information to share with other members of their team. The idea is that any individual participant would receive insufficient information on their own to solve the problem, so collaboration and communication is required. Scenarios are described in 68 short statements, typically a single sentence, called factoids.

Although the ELICIT platform holds considerable promise for use in future experimentation, it is unclear how systematically the four existing ELICIT scenarios have been designed. This is potentially problematic because achieving a high level of experimental control in future research would require being able to either exert control over dimensions such as scenario difficulty, or at least to understand the dimensionality of the existing scenarios. Toward this goal, this project has aimed to use the ELICIT scenarios as a starting point for developing a formal relational/structural framework that would enable the creation of systematically generating the content required for complex team scenarios. If this knowledge could be attained, this could be instructive for designing a workable approach to use for tightly controlled future scenario generation. All of these existing scenarios were analysed in some detail.

Analyses conducted in this course of this project slowly progressed from decomposition of the factoids into bits of information, to the level of factoids as a whole, to the level of the scenario as a whole. Through these explorations, some analyses emerged as the most helpful and promising way to think about how to dimensionalize each level.

A number of key observations were made through this process. The primary observation is that while the scenarios have an underlying structure which allows both deductive reasoning and inference, there seems to be little ability for the systematic creation of factoids or manipulation of scenario difficulty within the current generation system of ELICIT scenarios. Through the analyses conducted there appears to be a basic scenario structure upon which a framework for the understanding of difficulty, complexity, and other potentially interesting experimental manipulations could be built. The analyses of this project represent a start toward the goal of better understanding (and hence controlling) scenario characteristics and systematic scenario creation.

Sommaire

Élaboration d'un cadre de génération de contenu pour les scénarios d'équipe

Andrew Morton and Barbara D. Adams, Human*systems*® Incorporated; DRDC Toronto No. CR 2010-105; Defence R&D Canada – Toronto; March 2010

L'étude de la dynamique de la performance collective en laboratoire exige une plateforme expérimentale. ELICIT (*Experimental Laboratory for the Investigation of Collaboration, Information-sharing, and Trust,* de Parity Communications Inc., 2008) est l'une de ces plateformes. La plateforme ELICIT a été mise au point dans le but d'étudier expérimentalement l'organisation du commandement et contrôle (C2), les caractéristiques individuelles et d'équipe ainsi que la chaîne de valeurs centrée sur le réseau. ELICIT est un logiciel web qui est fondé sur le scénario général qu'une attaque terroriste se produira éventuellement dans l'avenir dans les limites d'une situation hypothétique et des équipes pouvant regrouper jusqu'à 17 membres doivent collaborer pour déterminer les particularités (c.-à-d. répondre aux question qui, comment, où et quand) de l'attaque en cours. À cette fin, chaque membre de "l'équipe" reçoit des unités d'information distinctes et il doit décider quel renseignement il partagera avec les autres membres de son équipe. Le principe est qu'aucun participant ne recevra suffisamment de renseignements pour résoudre le problème, il faut donc avoir recours à la collaboration et à la communication. Les scénarios sont décrits en 68 courts énoncés, habituellement formulés en une seule phrase, appelés factoïdes.

Bien que l'utilisation de la plateforme ELICIT soit prometteuse pour des expériences futures, on ne sait pas exactement quelle approche systématique a été utilisée dans la conception des quatre scénarios ELICIT existants. Cette situation peut potentiellement constituer un problème car l'atteinte d'un haut niveau de contrôle expérimental lors des recherches futures pourrait exiger de pouvoir exercer un contrôle sur des dimensions telles que la difficulté du scénario ou au moins sur la capacité de comprendre la dimensionnalité des scénarios existants. Dans ce but, le présent projet visait à utiliser les scénarios ELICIT comme point de départ à la mise au point d'un cadre relationnel/structurel formel qui permettrait la création systématiquement de contenu qui serait nécessaire pour les scénarios d'équipes complexes. Une fois obtenue, cette connaissance pourrait être instructive lors de la conception d'une approche pratique pouvant être utilisée pour la création de scénarios futurs strictement contrôlés. Tous les scénarios existants ont été analysés avec une certaine attention.

Les analyses menées dans le cours de ce projet ont progressé de la décomposition des factoïdes en segments d'information jusqu'au niveau des factoïdes dans leur ensemble, puis au niveau du scénario dans son ensemble. Grâce à ces explorations, certaines analyses se sont avérées être des voies plus utiles et prometteuses de concevoir la façon de dimensionner chaque niveau.

Un certain nombre d'observations clés ont été faites au cours de ce processus. La principale observation touche le fait que bien que les scénarios possèdent une structure sous-jacente qui permet de faire à la fois des raisonnements déductifs et de l'inférence, la création systématique de factoïdes et la manipulation du niveau de difficulté des scénarios semblent difficiles à l'intérieur de l'actuel système de génération de scénarios ELICIT. Les analyses effectuées ont permis de dégager ce qui semble être une structure de scénario de base à partir de laquelle il serait possible de bâtir un cadre permettant de comprendre la difficulté et la complexité des scénarios ainsi que d'effectuer d'autres

manipulations expérimentales potentiellement intéressantes. Les analyses de ce projet constituent un point de départ vers une meilleure compréhension (et ainsi un meilleur contrôle) des caractéristiques d'un scénario et de la création systématique de scénarios.

This page intentionally left blank.

Table of Contents

List of Figures

List of Tables

1. Background

Experimental research in team performance dynamics requires a platform from which to base the study. One such platform is ELICIT, the Experimental Laboratory for the Investigation of Collaboration, Information-sharing, and Trust (Parity Communications Inc., 2008). The Command and Control Research Program (CCRP) sponsored the development of ELICIT for information, cognitive, and social domain experimentation. ELICIT can be used to experimentally study command and control (C2) organization, team and individual characteristics, and the network-centric value chain. ELICIT is an instrumented platform from which the effect of experimental manipulations can be logged and recorded. The current version of ELICIT is a web-based software that can accommodate up to 17 participants at one time. The overall scenario is that a terrorist attack will occur at some point in the future within a hypothetical situation, and participants must work together to determine the details of the pending attack. To do this, each member of the "team" receives distinct units of information, and they must choose what information to share with other members of their team. The idea is that any individual participant would receive insufficient information on their own to solve the problem, so collaboration and communication is required. Successfully solving the scenario requires the team to discern the correct answer to each of four questions: Who, What, Where, and When. Colour names are used as possible actors (e.g. Red Group), institutions, infrastructure and individuals are used as possible targets (e.g. embassy), Greek letters are used as country names (e.g. Psiland), and timing include month, day, hour and AM/PM.

Scenarios are described in 68 short statements, typically a single sentence, called factoids. Of these 68 factoids, 4 are categorized as expert (E), 13 are key (K), 17 are support (S), and 34 are noise (N). The 68 factoids are assigned to the four questions (e.g., Who, What) evenly, with even numbers of E, K, S, and N factoids per question as best possible. In the standard 17 participant experiment, each participant receives one E or K factoid, 1 S factoid, and 2 N factoids. Factoids are presented to participants in three waves, typically separated by 5 minute intervals, with two factoids presented in the first wave and one in each subsequent wave. As participants are presented with factoids, they can choose whether or not to share the factoid with their fellow participants. The C2 organization being studied will dictate the rules surrounding the sharing of information. Each factoid is coded with a unique index number (1 to 68), distribution wave (1, 2, or 3), type (question), impact (E, K, S, or N), and count of the factoid within its type-impact category (1 to 9). Four different scenarios have been developed by the designers of ELICIT for use in experimentation, with the potential for additional scenarios to be used.

Although the ELICIT platform seems to hold considerable promise for use in future experimentation, it is unclear how systematically the current scenarios have been designed. This is potentially problematic because using the scenarios for experimentation would require more understanding of how the ELICIT scenarios have been designed, in order to achieve a high level of experimental control in future research. Toward this goal, this project has aimed to use the ELICIT scenarios as a starting point for developing a formal relational/structural framework that would enable the creation of systematically generating the content required for complex team scenarios. If this knowledge could be attained, this could be instructive for designing a workable approach to use for tightly controlled future scenario generation.

The chapters that follow describe our approach to achieving this goal.

This page intentionally left blank.

2. Analysis Descriptions

Despite the exploratory nature of this work, it was important to take a systematic approach to attempting to analyze the content of the existing ELICIT scenarios, with a view to understanding their form and structure. The following sections provide narrative descriptions of the analyses conducted in roughly chronological order. For each analysis a summary description with rationale, highlights, and lessons learned are presented. Further details are provided in the accompanying annexes.

2.1 General Scenario Observations

Through immersion in the scenarios during the analyses, a number of general observations were made regarding the scenarios' structures and coding. Scenarios 1 and 2 are highly similar, using the same logical sequences, merely changing names of actors, targets, countries, and timings. Scenario 3 is similar to 1 and 2 but with some significant structural differences. The structural differences are primarily demonstrated in a greater dependence on interim conclusions regarding the answers to What and Where. Scenario 4 is substantially different from the other three scenarios in structure.

Recall that the scenario developers classify each factoid, in terms of whether it was Expert, Supporting, or Noise. We first explored these classifications. Our observations were that Expert and Key classifications were generally accurate in labelling factoids which contained crucial information to the solving of the scenario. Surprisingly, however, classification of Support and Noise factoids generally had showed less distinction among them. While Supporting factoids were more often related to excitation of the answer and inhibition of the non-answer, occasionally factoids with no useful information would be classified as Support by the scenario developers. Noise factoids were almost always irrelevant to actually solving the scenario; some noise factoids did in fact contain useful, although not critical, information toward solving the scenario.

The type classification (Who, What, Where, When) of factoids provided by the scenario developers were also generally accurate for Expert and Key factoids but much more variable for Support and Noise factoids. Type classification is often complicated by single factoids linking components of multiple questions. Despite this complication, many Support and Noise factoids are type classified incorrectly, and are purported to relate to questions that they do not actually address. This erroneous type classification may have been motivated by the need to ensure equal distribution of factoids by type, but this is speculative. The scenario developers indicated in their documentation that some questions are more complicated and hence may have more factoids devoted to them.

In all scenarios provided, the correct answer can be logically determined through deductive reasoning with the use of Expert and Key factoids. Support factoids generally assist in determining the correct answer to the scenario through inferential reasoning. Noise factoids generally do not offer any insight into the correct answer to the scenarios. In general, Who is determined using a process of identifying viable alternatives, elimination of all but one alternative, and accumulation of confirmatory evidence. The What and Where questions were often inter-related in the sample scenarios, such that solving of one would require at least partial resolution of the other. Solving the What and Where use a similar process as solving Who of identification, elimination, and confirmation. To solve When each part must be identified and confirmed, generally with no non-answer alternatives to eliminate. Confirmation of When parts is often accomplished through the Who solution.

2.2   Factor Inventory & Categorization

The first analysis conducted examined scenario 1, during which the subjects or factors (e.g. possible colour groups to answer the Who question) were inventoried and categorized. See Figure 1 for the full categorization of all scenario 1 factors.
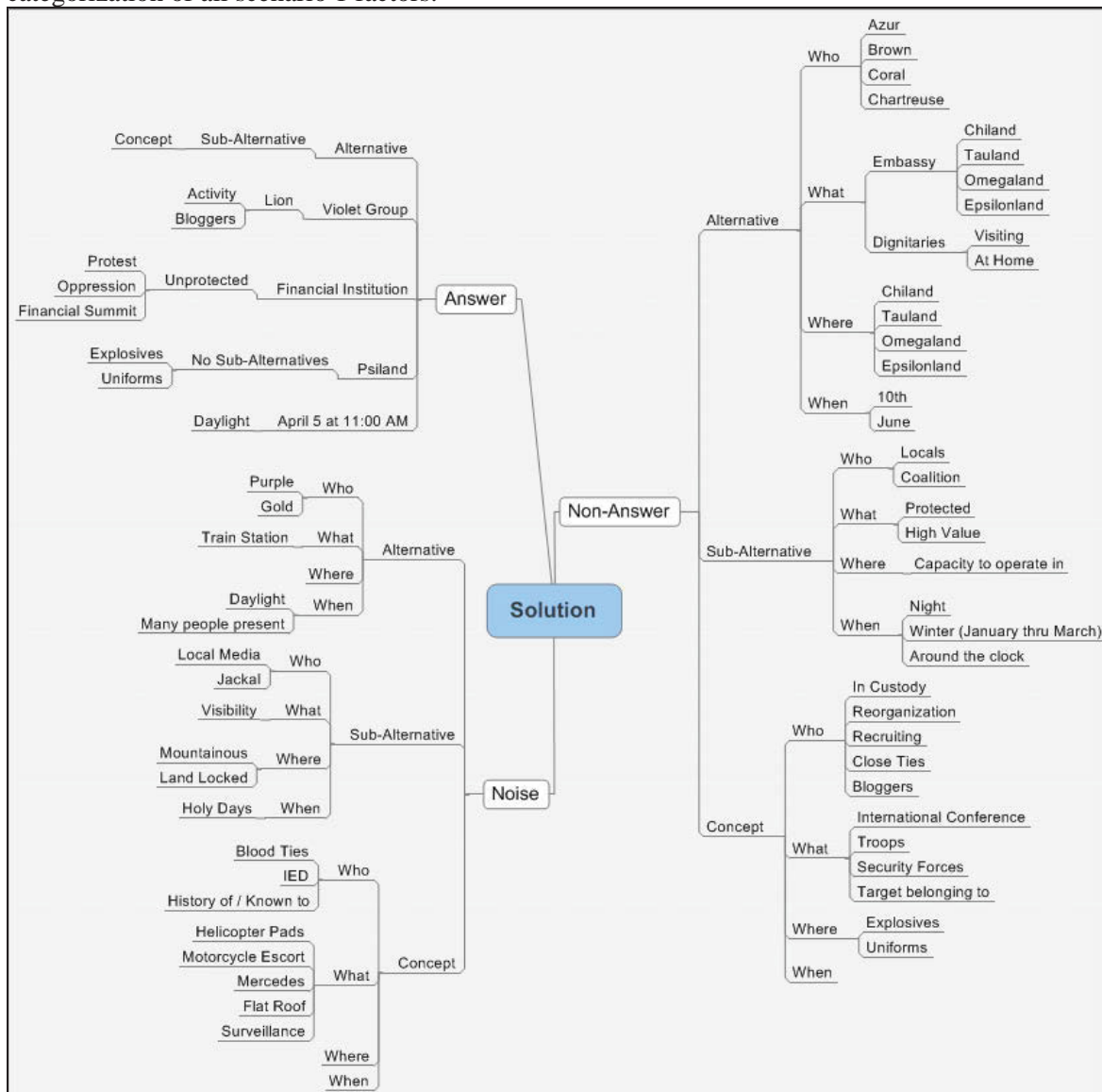
Figure 1: Scenario 1 factor categorization

The ELICIT framework uses colour named groups for Who, the target is What, Greek letter land countries for Where, and month, day, hour, and AM/PM for When. The inventory simply captured all of the alternatives to each of these questions mentioned within the factoid set. This analysis showed additional factors within the factoid sets that, while related to one of the four questions, were not direct answers to any question. For example, the Lion is mentioned as an involved actor but is not associated with a colour group. These factors were termed sub-alternatives and inventoried. Other factors were also discovered that implied or suggested a particular sub-alternative; factors falling into

this category were termed "concepts". For example while protection was a sub-alternative of the target, security forces were a concept used to imply protection. Factors were also classified as being components of the answer, the non-answer, or noise. Non-answer components were viable to be considered as the answer (i.e. mentioned in Expert, Key, and/or Supporting factoids), while noise components were only introduced in Noise factoids.

The factor inventory approach was used with all four scenarios provided in the ELICIT software. Alternatives and sub-alternatives which were only introduced in Noise factoids were excluded from the inventory. Concepts were not inventoried. Given these categorizations, it was then possible to explore the frequencies of each of the different types of factors. Analyses were also conducted to identify the number of factoids needed to deduce the correct answer, a result of the solving matrices analysis to be described in 2.6 of this report. Finally, the relative difficulty of each scenario as described by the scenario developers is included in the column headers. All of this information for each of the 4 scenarios is shown in Table 1.[1]

Table 1: Factors and solution factoid requirements for four sample scenarios

Questions \ Scenarios		1 (middle)	2 (hardest)	3 (hardest)	4 (easiest)
Who	Alternatives	5	5	5	7
	Sub-Alternatives	2	2	3	1
	Factoids Needed	5	5	8	6
What	Alternatives	7(8)	7	3(5)	3
	Sub-Alternatives	1	1	2	1
	Factoids Needed	4	5	4	4
Where	Alternatives	5	5	5	5
	Sub-Alternatives	2	2	2	1
	Factoids Needed	7	4	7	12
When	Alternatives	4 parts	4 parts	2 parts	4 parts
	Sub-Alternatives	-	-	-	-
	Factoids Needed	9	9	8	8

This analysis was a useful first step in understanding the scenarios and the structural and relational aspects of them. The factor categorization provided terminology to for the discussion of the scenarios. Furthermore the factor inventory demonstrated that manipulations of difficulty are not simple changes of in scale of the problem space.

2.3 General Solution Process

The typical process for deducing the solution from the Expert and Key factoids uses definitive excitation of one or more sub-alternatives, definitive inhibition of some alternatives and/or sub-

---

[1] Note that the number of What alternatives in scenarios 1 and 3 can be interpreted as two different totals and hence the two numbers are given for each. Scenario 3 combines month with day and hour with AM/PM to reduce the When question from 4 parts to 2 parts.

alternatives, and the linking and delinking of alternatives and sub-alternatives. One or two Key factoids from each scenario provide only existence information, laying out the viable alternatives to one or two questions.

The supporting factoids consisted of statements giving probabilistic excitation and linking of the answer alternatives/sub-alternatives, probabilistic inhibition and linking of non-answer alternatives/sub-alternatives, delinking of answer and non-answer alternatives/sub-alternatives, and establishing existence of alternatives/sub-alternatives. Supporting factoids would often contain additional information that could not be considered supporting (e.g. probabilistically linking a non-answer alternative to an answer sub-alternative); however taken as a group, the supporting factoids performed the functions listed above.

Noise factoids provided distraction and/or detraction from the scenario solution by introducing new alternatives and sub-alternatives that should not be considered viable answers, linking and delinking alternatives and sub-alternatives in non-meaningful ways, and using unclear or ambiguous wording that would require uncertain interpretation. Noise factoids more frequently established existence of noise components and performed linking and delinking components, and less frequently related to excitation and inhibition statements. Noise factoids would never definitively excite any components (answer or non-answer), definitively inhibit the answer components, or definitively inhibit viable non-answer components.

2.4   Decomposition & Categorization Analysis

One of the first analyses conducted on the ELICIT scenarios was to decompose the factoids into information bits and categorize these bits by salient characteristics. All factoids of scenario 1 were broken into a series of information bits. Information bits were the smallest meaningful chunks of data that could be extracted from the factoid (e.g. Tauland is land locked). These decomposition analyses were conducted in isolation from each other and from the final answer, such that both meaningful and irrelevant information bits were distilled. The assumption was made that participants knew the naming conventions of Greek letters for countries and colours for aggressor groups. Existence information was generally not included (e.g. There is a Tauland), with the exception of interaction of targets in countries (e.g. Tauland has an embassy in Epsilonland). From analyses conducted latter in the project it was understood that existence information is valuable; however, this analysis was not revisited as more fundamental limitations with this approach had been discovered. Excitatory information bits (that is, bits that appeared to support the subject as the conclusion) were scored as positive one, inhibitory information bits (bits that opposed the subject as the conclusion) were scored as negative one, and existence information was scored as zero (bits that neither supported nor opposed the subject as the conclusion).

Information bits were then categorized as definitive or probabilistic. Definitive bits were defined as statements that are authoritative, specific, and precise. Probabilistic bits were defined as statements that are suggestive, general, and imprecise. Within the definitive classification, bits were further subdivided as singular, disjunctive, conjunctive, sub-alternative, and/or conditional. Within the probabilistic classification, bits were further subdivided as singular, multi, interactive, and/or sub-alternative, as shown in Table 2.

Table 2: Information bit definitions

Type and definition	Example
Definitive singular were single alternative/sub-alternative information bits;	e.g. "The Lion is involved".
Definitive disjunctive were multiple alternative/sub-alternative information bits separated by disjunction (or);	e.g. "The target is a coalition member embassy, visiting dignitary, or financial institution".
Definitive conjunctive were multiple alternative/sub-alternative information bits separated by conjunction (and);	e.g. "All high value targets belonging to Tauland and Epsilonland are well protected".
Definitive sub-alternative were statements referencing sub-alternatives;	e.g. "The Lion will not risk working with locals".
Definitive conditional were multiple alternative/sub-alternative information bits conditional on other information bits being true;	e.g. "The Lion doesn't operate in Chiland".
Probabilistic singular were single alternative/sub-alternative information bits;	e.g. "There has been an increase in messages intercepted in Psiland".
Probabilistic multiple were multiple alternative/sub-alternative information bits;	e.g. "The Azur, Brown, Coral, Violet, or Chartreuse groups may be planning an attack".
Probabilistic interactive were multiple alternative/sub-alternative information bits linking multiple questions	e.g. "The Azur, Brown, Coral, and Violet groups have the capacity to operate in Tau, Epsilon, Chi, Psi and Omega-lands".
Probabilistic sub-alternative were statements referencing sub-alternatives;	e.g. "The Violet group prefers to operate in daylight".

Further analysis at the factoid level was simultaneously conducted. Factoids were scored for the number of questions addressed, number of alternatives referenced, number of sub-alternatives referenced, and number of info bits extracted. Factoids that provide information to exhaustively include or exclude an alternative or sub-alternative (all, only, etc.) were scored as independently exclusive; e.g. "All high value targets belonging to Tauland and Epsilonland are well protected". Expert and Key factoids were also categorized as "Identify" the correct answer alternatives or "Eliminate" incorrect answer alternatives and "Direct" referencing alternatives, "Indirect" referencing sub-alternatives, or "Link" linking alternatives and sub-alternatives.

Results of these decomposition, categorization, and scoring analysis were described in frequency statistics. Differences between Expert and Key factoids, Supporting factoids, and Noise factoids were examined. Graphical analysis of the results of this analysis provided no overwhelming findings but provided further insight into the structural aspects of scenario 1. From this analysis it became evident that Expert and Key factoids were primarily composed of definitive information bits, Supporting factoids were primarily composed of probabilistic information bits, while Noise factoids distributed between definitive and probabilistic information bits - see Figure 2. This finding influenced the subsequent analyses that were conducted.
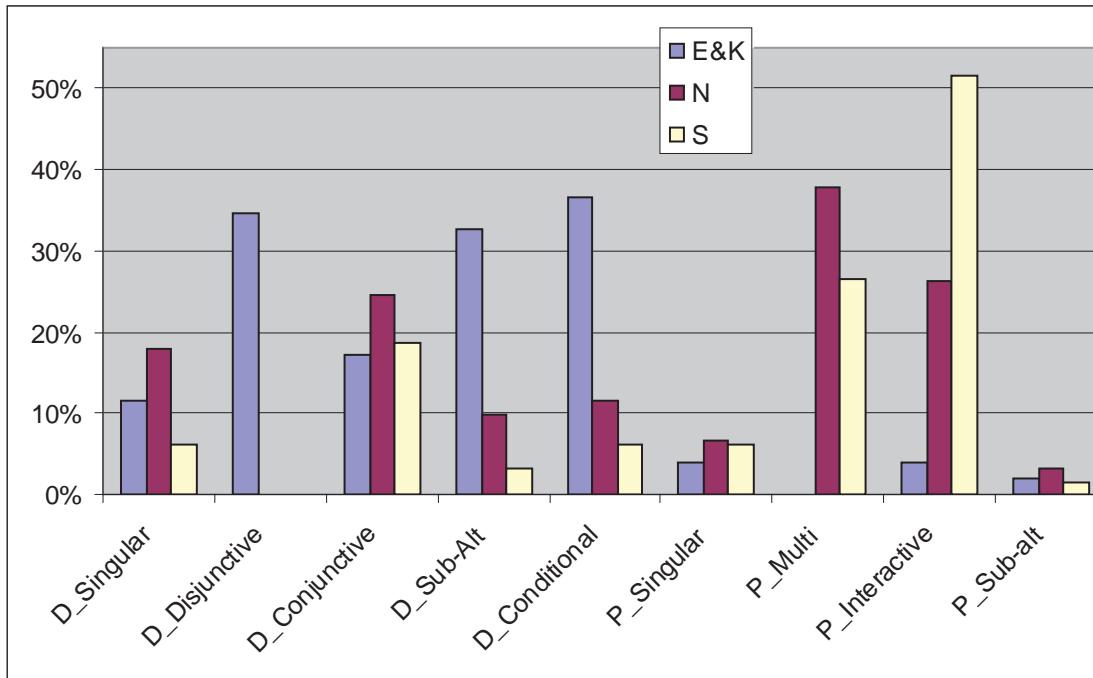
Figure 2: Percentage distribution of information bits by factoid impact category

A significant limitation of this analysis was revealed in that individual information bits may be classified in more than one of the categories defined. Furthermore, the relevance of the information bit to the final answer was not considered; therefore, although this allowed determination of the structural characteristics of the bit, the value of the bit could not be determined. As a result, pursuit of this analysis was discontinued; however, through this first attempt invaluable familiarity with the scenario structure and details was gained. See Annex A for further details of this analysis.

2.4.1 Information Bit Re-Categorization

Information bits were then re-analyzed using a new set of categories. Instead of focusing solely on the structural characteristics of each information bit and factoid, information bits were categorized for their potential value in light of the certainty of the information given. After slight modification to the listing of information bits from lessons learned in the first analysis, the following categories were used: truth, worth, and definitive/probabilistic. For the "truth" category, information bits were classified as being inherently true (i.e. internally verifiable) or conditional (i.e. truth is dependent on other information). The "worth" category divided information bits into those adding value to determining the answer and those whose value is conditional on other information. Finally, information bits were classified as being definitive or probabilistic. The percentage distributions of information bits by factoid type for each of these categories are given in Table 3.

Table 3: Categorization re-analysis percentage distributions

Impact	Truth		Worth		Definitive/Probabilistic	
	True	Conditional	Value	Conditional	Definitive	Probabilistic
Expert	100%	0%	50%	50%	100%	0%
Key	60%	40%	3%	97%	59%	41%
Support	93%	7%	20%	80%	24%	76%
Noise	75%	25%	4%	96%	46%	54%

The definitive/probabilistic classification appears to be the most distinctive categorization for the different impact factoids, where Expert factoids contain definitive information bits, Key factoids are mixed definitive and probabilistic, Supporting factoids are also mixed but with a greater percent probabilistic bits, and Noise factoids are relatively evenly mixed.

2.5 Experimental Manipulations

As the long term goal of this project is to assist future experimentation efforts, we also worked to identify possible ways to dimensionalize the characteristics of the scenarios. This was done with a view to being able to design future scenarios (or use existing scenarios) while varying the desired dimension. Some potential experimental manipulations are shown in Figure 3.
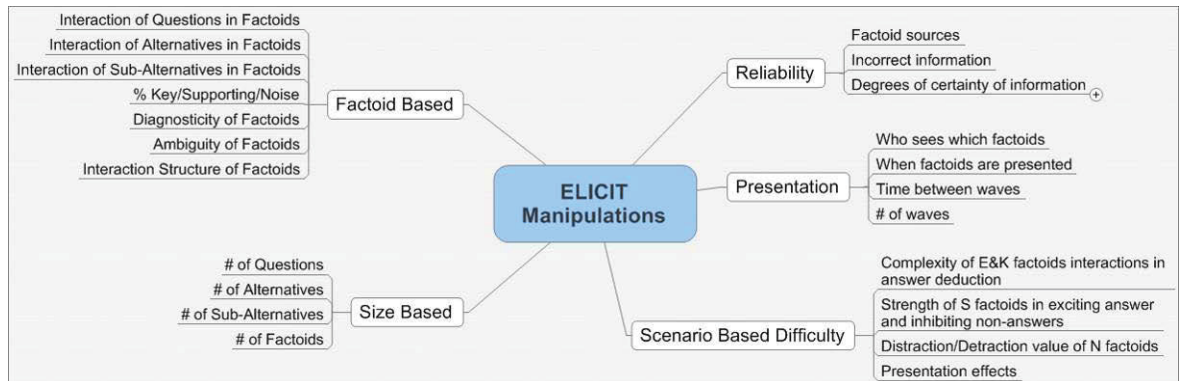
Figure 3: ELICIT experimental manipulations

Five broad categories were defined, with possible overlap between categories: factoid based, sized based, reliability, presentation, and scenario based difficulty. Factoid based manipulations seek to alter the difficulty or complexity of the scenario through changes to the structure, composition, and interaction of the factoids. Size based manipulations are simply changes to the number of factors within the scenarios. Reliability manipulations vary the certainty of the information comprising the scenario. Presentation manipulations are strictly relating to the distribution of factoids to the participants during the ELICIT scenario. Scenario based difficulty manipulations stem from the four components theory of scenario difficulty discussed below, where a study could examine the impact of changes within each of the four components.

2.6 Solving Matrices

Following the decomposition and categorization analysis, a different approach was taken in which the solution to the scenario would play a central role. Isolating the E and K factoids, each question of the scenario was examined to determine which factoids were needed to deduce the correct answer. For some questions in some scenarios, the factoid requirements to eliminate each non-answer alternative

could be simply listed; however, for other questions the interaction of alternatives across questions prohibited the deduction of a single question's answer in isolation. This examination led naturally to the creation of "solving matrices". Solving matrices plot the viable alternatives of one or more questions in the different axes of a table. Factoids and combinations of factoids are then indicated in each cell to exclude each alternative or alternative combination. In this way, the deduction of a solution with question interaction can be clearly represented. An example solving matrix for the questions of What and Where in scenario 1 is presented below (see Table 4) - for complete solving matrices to all 4 scenarios see Annex A. Note that where multiple groups of factoid numbers are listed in a cell separated by a comma, either combination can be used to exclude the alternative combination.

Table 4: Solving matrix of What and Where for scenario 1

		What						
		Dignitaries	Tauland Embassy	Epsilonland Embassy	Chiland Embassy	Psiland Embassy	Omegaland Embassy	Financial Institution
Where	Tauland	2&29		2&22	2&39	2&39	2&39	2&22
	Epsilonland	2&29	2&22		2&39	2&39	2&39	2&22
	Chiland	1&3, 2&29	1&3, 2&22	1&3, 2&22		1&3, 2&39	1&3, 2&39	1&3
	Psiland	2&29	2&22	2&22	2&39		2&39	
	Omegaland	2&29, 2&47	2&22, 2&47	2&22, 2&47	2&39, 2&47	2&39, 2&47		2&47

The solving matrix technique allowed the clear presentation of which factoids were needed to resolve each question of a scenario. From this, the number of factoids required for each question's resolution in each of the 4 scenarios was determined and is presented in Table 1. While there is no obvious pattern in the number of factoids required to solve the problem compared to the relative difficulty of scenarios according to ELICIT designers, there is variability in factoids required between questions and between scenarios.

2.6.1   Factoid Contribution Metrics

Following the development of the solving matrices technique, a means by which to quantify the contribution of each E&K factoid to the solution was developed. A simple count of the number of times each factoid was used throughout the solving matrices was a first pass; however this approach lacked the precision to differentiate most factoids. A metric of value (V) was then developed which captured the number of alternatives in which the factoid was used to exhaustively exclude the question, with fractional contribution for partial exclusion (i.e. only select alternatives of another question could be excluded). A deficiency with this metric was immediately identified in that it does not account for the conditionality of factoids, in which multiple factoids may be needed in combination in order to conclusively exclude a matrix cell. As a result, a weighted value metric (W) was developed which would use the same basis as V but be fractionally weighted for the conditionality of a factoid.

An example solving matrix is given for the interaction of two questions in Table 5.

Table 5: Example solving matrix

		Question 2		
		Alternative X	Alternative Y	Alternative Z
Question 1	Alternative A	1&2	1&5	7
	Alternative B	2&3	2&6	2&8
	Alternative C	2&4	 	7

A simple count of factoid 2 usage would give a score of 5. Factoid 2 is used in exclusion of all question 1 alternatives of alternative X and alternative B of both alternatives Y and Z, the V metric of factoid 2 for question 2 would be 1.67 (3/3 for Alternative X plus 1/3 for Alternative Y plus 1/3 for Alternative Z). In the same example, the W metric of factoid 2 for question 2 would be 0.88 (1/2 of 3/3 = 0.5 for Alternative X plus 1/2 of 1/3 = 0.167 for Alternative Y plus 0.167 for Alternative Z). The W metric appears to be a promising means by which to quantify the relative contribution of E&K factoids to the scenario solution.

2.7 Framework

A conceptual framework of information bits was built to reflect the analyses conducted to this point. The framework was represented by a three dimensional cube, with axes of Proximity, Outcome, and Strength, as shown in Figure 4.
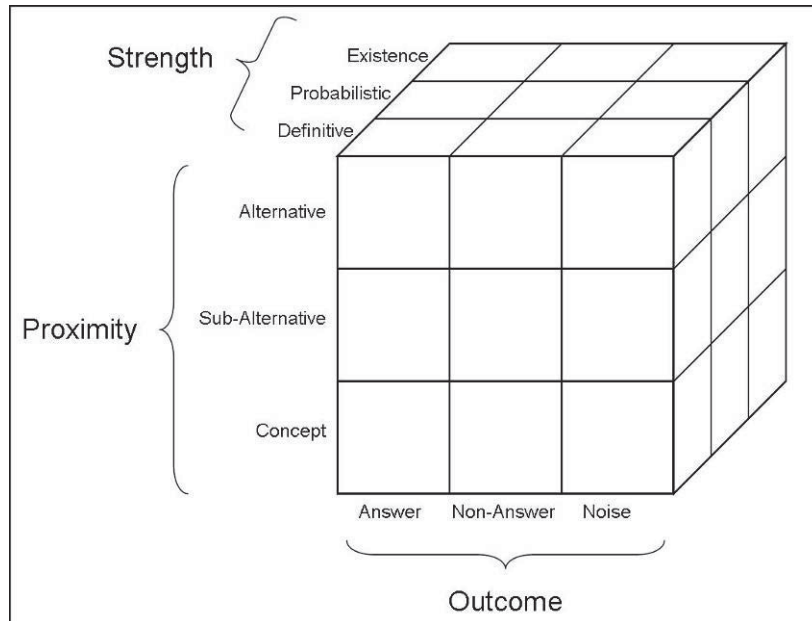
Figure 4: Conceptual framework of information bits

Proximity referred to three layers of components being alternatives, sub-alternatives, and concepts. It was theorized that alternatives are more proximal components to the answer and non-answers, whereas sub-alternatives must be linked or delinked from alternatives to influence the scenario solutions. Concepts are used to imply sub-alternatives.

Outcome referred to three classifications of components: answer, non-answers, and noise. The scenario solution is comprised of the answer components. Non-answer components were defined as

viable alternatives to the answer components, whereas noise components were non-viable components. The distinction between non-answer and noise components was fuzzy, with the clearest indication being that non-answer components would be established and discounted in the E&K factoids, while noise components would be introduced as distraction in the N factoids. It was theorized that answer components would have greater value than non-answer components, which in turn were of greater value than noise components.

Strength of the information had three levels, in order of hypothesized value being definitive, probabilistic, and existence. Definitive information bits provide an authoritative excitation, inhibition, linking, or delinking of components. Probabilistic information bits give possible excitation, inhibition, linking, or delinking of components. Existence information bits establish the presence of components.

Positive value was theorized to be created through excitation of answer components, inhibition of non-answer components, linking within answer and non-answer categories, and delinking between answer and non-answer categories. Conversely, opposite actions to those theorized as having positive value were given negative value. Finally, actions not considered to have either positive or negative effects would be considered neutral noise.

2.7.1 Scoring

A scoring scale was then devised for the framework such that following factoid decomposition, each information bit could be scored for its relative value. The value of all information bits contained within a factoid could then be summed to determine the relative value of a factoid. For the scoring scale, a minimal point was established at existence noise concepts and given a value of 1. A maximal value was established at the intersection of definitive answer alternatives and given a value of 27. The scale was then evenly graduated between the two origins to give scoring value to all levels of the framework - see Figure 5.
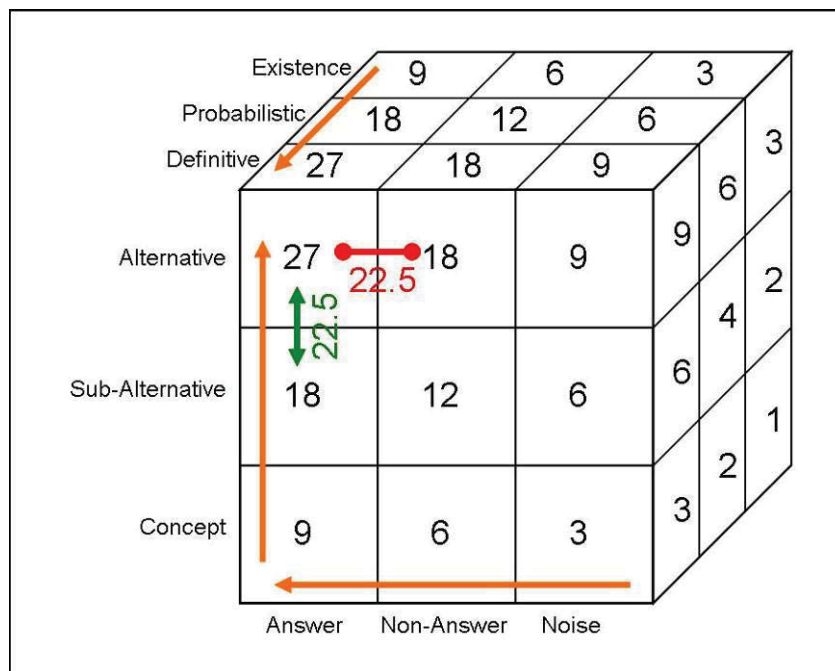

Figure 5: Framework scoring first draft

Links and delinks were valued as the average between the two cells housing the components being linked or delinked. Finally a positive sign was assigned to actions previously hypothesized as adding positive value and negative sign was applied to the converse actions. It was immediately apparent that this scoring system gave too much weighting to noise and was problematic in dealing with links and delinks. As a result, a revised scoring scale was developed.

2.7.2   Revised Scoring

The revised scoring scale used multiplication weightings for each of the three axes. Outcome was valued as 0 for noise, 1 for non-answer, and 2 for answer. Proximity was valued as 1 for concepts, 2 for sub-alternatives, and 3 for alternatives. The strength factor was scaled as 1 for existence, 2 for probabilistic, and 3 for definitive. The multiplication of the weightings from each axes provided the value assigned to each cell. In order to account for linking and delinking actions, information bit values were determined by multiplying the value of individual components together. In the case of excitation or inhibition the value of the individual component was squared. Again a positive sign was assigned to actions previously hypothesized as adding positive value and negative sign was applied to the converse actions. The revised scoring system was applied to the information bit decomposition for E&K factoids. Component information bit scores for each factoid were summed to give a factoid valuation score.

2.8   Correlations Between Analyses

The factoid valuations derived from the revised framework scoring of component information bits were then compared to the factoid valuation metrics from the solving matrices analyses (count, V, and W). Simple correlations indicate the strongest relationship between the revised framework scores and the W metric (0.758), with weaker correlations to V (0.698) and factoid use counts (0.395). This was an encouraging finding as it demonstrated convergence of the two analyses conducted at different levels of decomposition.

2.9   Supporting Matrices

In an attempt to capitalize on the success of the solving matrices analysis technique, matrices of S factoids were developed for Scenario 1. As with the solving matrices, S factoids were classified by the alternative to which they applied, with the What and Where questions again interdependent. For the supporting factoids, it was also necessary to classify the action of the factoid as "helping" or "hindering" determination of the correct solution for each matrix cell to which they would apply. For example, a S factoid may probabilistically excite the three Who alternatives, only one of which is part of the solution. In this case, the probabilistic excitation of the answer Who alternative would be considered "helping", while the probabilistic excitation of the non-answer Who alternatives would be considered "hindering". A simple count number of uses of each factoid was conducted. For each instance of "helping" a factoid would be given positive one, while for each instance of "hindering" a factoid would be given negative one. Thereby, net total helping-hindering were compiled - see Table 6.

Table 6: Supporting matrices counting measure

Factoid	Who	What	Where	When	"Helping" (✓)	"Hindering" (✗)	Net
5	✓ ✗✗✗✗				1	4	-3
6	✓ ✗				1	1	0
7	✓				1	0	1
12	✓				1	0	1
24			✓ ✗✗		1	2	-1
25		✓	✗✗✗		1	3	-2
26			✓ ✗✗		1	2	-1
27	✓ ✗✗	✓			2	2	0
28	✓ ✗	✓			2	1	1
37	✓ ✗✗✗		✓ ✗✗✗✗		2	7	-5
44		✓	✗		1	1	0
46	✓ ✗		✓		2	1	1
48			✓		1	0	1
57		✓ ✗		✓	2	1	1
60				✓	1	0	1
61	✓ ✗		✓ ✗✗		2	3	-1
63	✓			✓	2	0	2

The minimal, or even negative, net effect of support indicated by the supporting matrices helping-hindering counts was unexpected. It appears that the total effect of support factoids in accumulating evidence for the answer and against the non-answers is more complex than this simple counting analysis can account for.

2.10   Conjunctive Normative Form

Stating Scenario 1 in Conjunctive Normative Form (CNF) was attempted. Only E&K factoids were used as these factoids alone are necessary to deduce the solution; however, the conditionality contained within the E&K could not be accurately captured in CNF with the analyst's limited experience in CNF. As a result, this effort was soon abandoned.

2.11   Solution Maps

Following a meeting with the project Scientific Authority (SA) to brief him on the progress to date and the results of analyses conducted, it was recommended that further efforts be focused on studying the structure of factoid interaction in deriving the solution. In response to this focus, a series of solution maps depicting the E&K factoid interactions in solving Scenario 4 were developed. The first of these solution maps used the Mindjet MindManager Pro 6 software, which is a computer application that facilitates the creation of mind map diagrams. From a central topic, branches are extended with further branching as needed. Links between branching of any level can be added as needed; however the graphics created are inherently hierarchical in structure and appearance. Factoids of Scenario 4 were organized according to the question to which they most closely applied - see Figure 6.
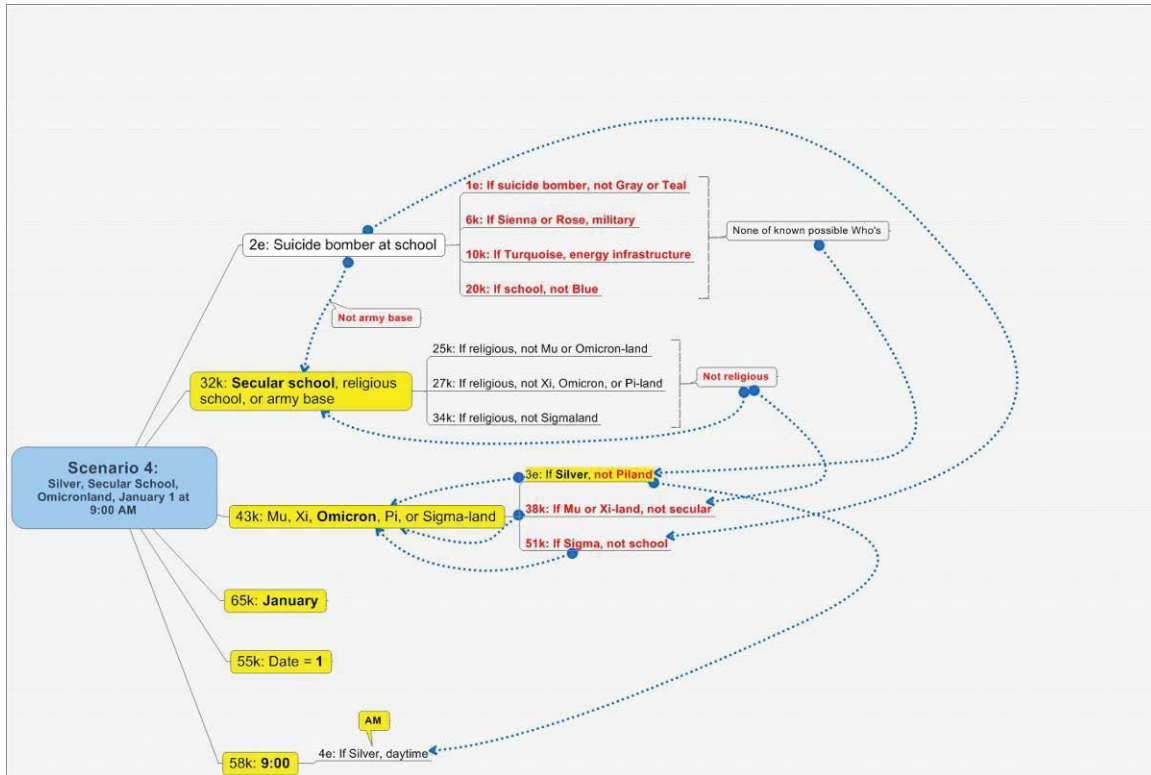
Figure 6: Mind map solution map by individual factoids

Each factoid was listed only once and links drawn as necessary to convey the dependence of factoids on each other. The text of each factoid was shortened and summarized to save space. Red text was used to indicate definitive elimination of alternatives with bold text and yellow highlights used to indicate the definitive identification of correct solution alternatives. Call-out boxes were added to indicate interim conclusions not explicitly stated in factoids. The resulting solution map depicted the logic pathways to determining the solution, but was messy in appearance with no salient characteristics, patterns, or message. As a result, a similar mapping strategy was attempted using a more flexible software graphics application.

The second solution map was developed using Microsoft Office Visio software and followed the same strategy as the first attempt. Visio is a diagramming program used for producing graphical diagrams, particularly suited to flow charts and network diagrams. This software allowed a more realistic process of knowledge accumulation allowing the determination of the correct solution, as shown in Figure 7,
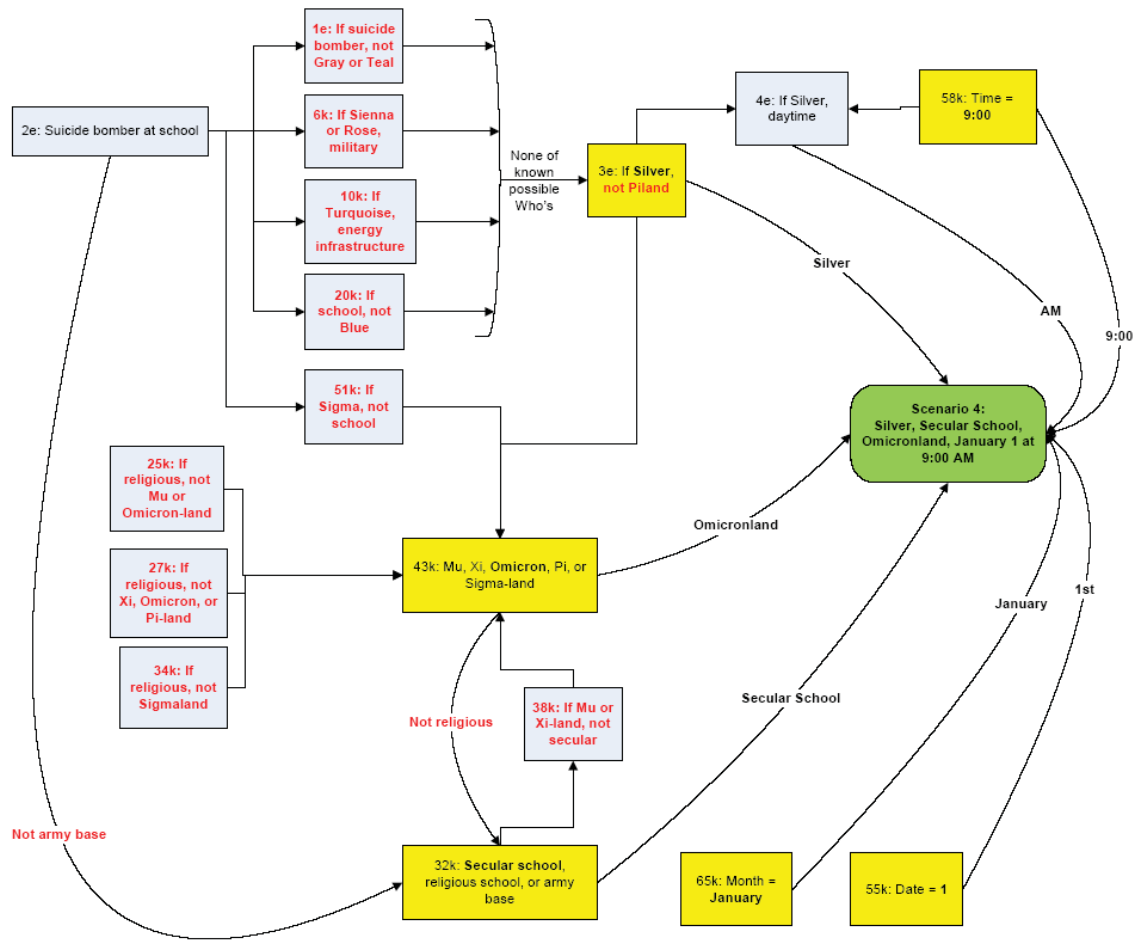
Figure 7: Flow chart solution map by individual factoids

Again shortened and summarized factoid text was used to save space. The same rules of factoid use and colour coding conventions as in the first attempt were used. While this second solution map more clearly presented the logical flow in determining the correct solution, it remained difficult to extract salient characteristics or patterns from the diagram. A third attempt at creating a simple, meaningful representation of factoid interaction again used Visio but focused on interim conclusions.

Where the first and second attempts at solution maps used individual factoids as nodes with relationships and links applied between them as appropriate, the third attempt used interim conclusions as nodes and all E&K factoids as lines adding information to the diagram or linking nodes, as shown in Figure 8.
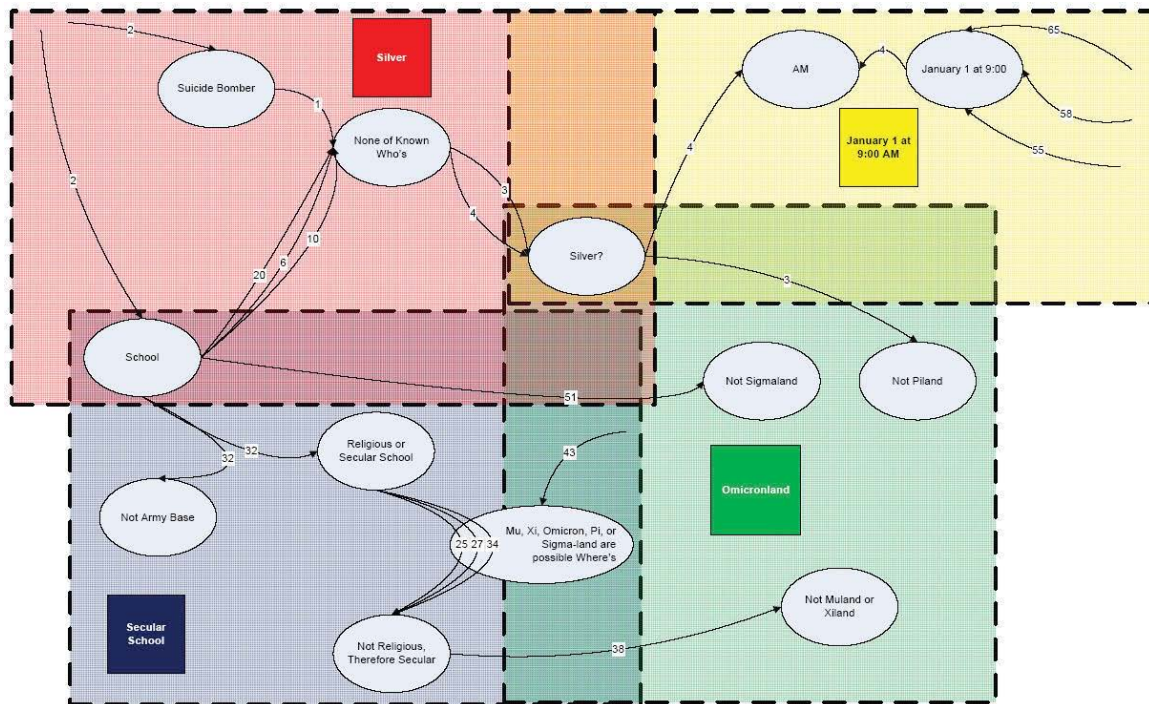
Figure 8: Solution map by interim conclusions

In this attempt, factoids were labelled by number only. The four questions were clustered in the four corners of the diagram (Who upper left, What lower left, Where, lower right, When upper right) and shaded boxes indicating the area of each question were added post-hoc. Several nodes were overlapped by multiple questions. Where possible each factoid was used only once; however, where multiple conclusions stemmed from a factoid, multiple lines for a single factoid were necessary. This mapping technique clearly clustered the questions and demonstrated how interim conclusions can be used in the determination of the answer to several different questions; however, the salient visual depiction of factoid interaction was still not apparent.

The fourth attempt at a solution map returned to the mind mapping software, but grouped factoids by question and alternative. The fourth solution map strategy laid out a base structure of each question, with all of the relevant alternatives branched within each question - see Figure 9. Factoids were re-used as many times as applicable, with only short summary text. Red text was used to denote factoids used in the definitive elimination of an alternative, with eliminated alternatives also indicated in red text. Where interim conclusions were needed across questions, hash line boxes indicated the boundaries of the interim conclusion and linking arrows pointed to the location the interim conclusion was needed. Supporting factoids were then added in black text and noise factoids in green text. While this solution map appears much more detailed (due to the inclusion of supporting and noise factoids), several important findings were drawn from this attempt. Starting with a base structure of questions, allowing the re-use of factoids wherever applicable, and stringing multiple factoids together in hierarchical relationships to represent interim conclusions appeared to be promising strategies to representing factoid interaction. From this fourth attempt at mapping the solution, the technique of solution trees evolved.

Figure 9: Solution map by question-alternative groupings
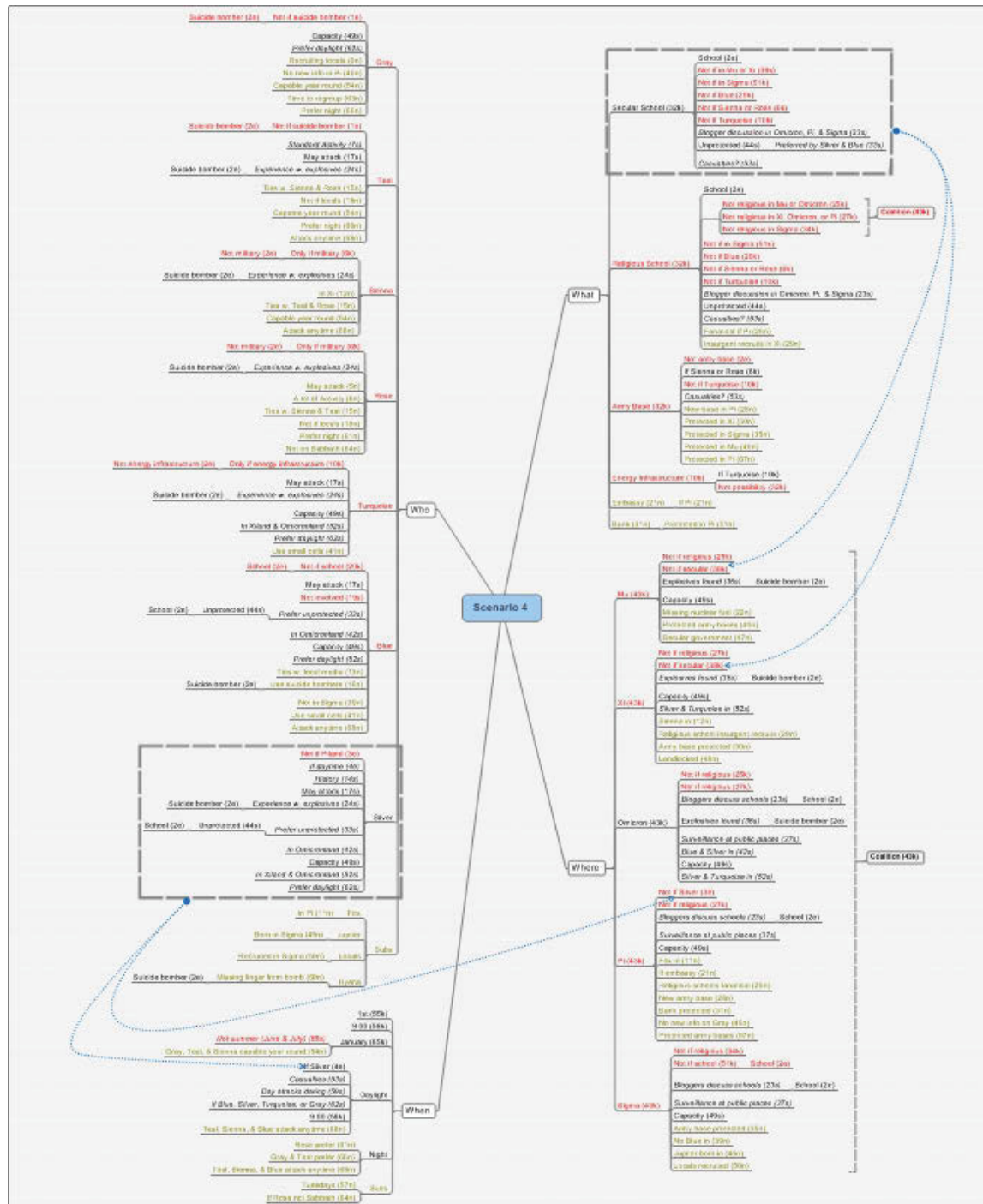
2.12 Solution Trees

Solution trees were the culmination of the solution mapping exercise intended to depict the logical interaction of factoids in determination of the correct solution. Drawing from the lessons learned in the four different solution mapping strategies, solution trees were designed to provide a visually salient representation of factoid interaction. Using a similar base structure to the fourth solution mapping technique with each question separated, E&K factoids were used as many times as necessary in hierarchical organization by conditionality to depict the logical sequence of determining the answer - see Figure 10 through Figure 13. Conditionality of factoids in terms of solution trees refers to aspects of a factoid that can only be used with further information. For example, factoid 8 of scenario 1 (The Lion is known to work only with the Azur, Brown, or Violet groups) is conditional in determining the answer to Who on factoid 1 (The Lion is involved) which establishes the three groups listed as the only viable actors, factoid 10 (All members of the Azur group are now in custody) which eliminates the Azur group, and factoid 14 (The Lion will not risk working with locals); factoid 14 is further conditional on factoid 13 (The Brown group is recruiting locals - intentions unknown) which eliminates the Brown group, leaving only the Violet group. Furthermore, factoids were reduced to the index number with no descriptive text. The resulting tree clearly illustrates the differences between scenarios in terms branching and length. Branching is a function of non-exclusivity of factoids, where multiple alternatives/sub-alternatives may be capable of satisfying (or necessary to resolve) the factoid statement (e.g. the three colour groups noted in factoid 8 of scenario 1). Length of branches in this depiction is a function of the conditionality of factoids, with a series of sequentially conditional factoids resulting in a longer branch. Hash lined outlines were used to visually highlight the repetition of a pattern of specific factoids, representing an interim or final conclusion. The solution trees of Scenarios 3 (Figure 12) and 4 (Figure 13) include the factoid impact classification while the solution trees of Scenarios 1 (Figure 10) and 2 (Figure 11) do not; however, factoids numbered 1 through 4 are always the Expert factoids and all other factoids used in the solution trees are Key factoids.

The scenario trees facilitated a rapid appreciation of the relative complexity of E&K factoid interaction in a scenario. The similarities between Scenarios 1 and 2 are explicit and remarkable when presented in solution tree form. The branching, length, and pattern repetition are nearly identical between these two scenarios; thereby validating previous descriptions of these two scenarios as highly similar.

From the solution trees of the four scenarios provided by ELICIT's developers, a number of general observations can be made. For each of the Who, What, and Where questions, a single factoid is used to establish the answer alternative and in most cases bound the possible non-answer alternatives as well. From this initial factoid, branches are then split off to eliminate each of the non-answer alternatives. The number of branches is determined by the number of non-answer alternatives and the exclusivity of the factoids. For the question When, four branches are immediately established corresponding to the four parts of the question (month, day, hour, AM/PM) with the exception of Scenario 3 which combines month with day and hour with AM/PM. The length and subsequent branching of each non-answer elimination branch is a function of the conditionality of factoids. The factoids at the tip of each branch are definitive and unconditional, and are a limited subset of the scenario's factoids. Often an interim conclusion, or even final answer to one question, will be utilized in other areas of the solution tree to resolve the conditionality of factoid thereby increasing the complexity of factoid interactions and size of the solution tree without changing the total number of unique factoids (17).
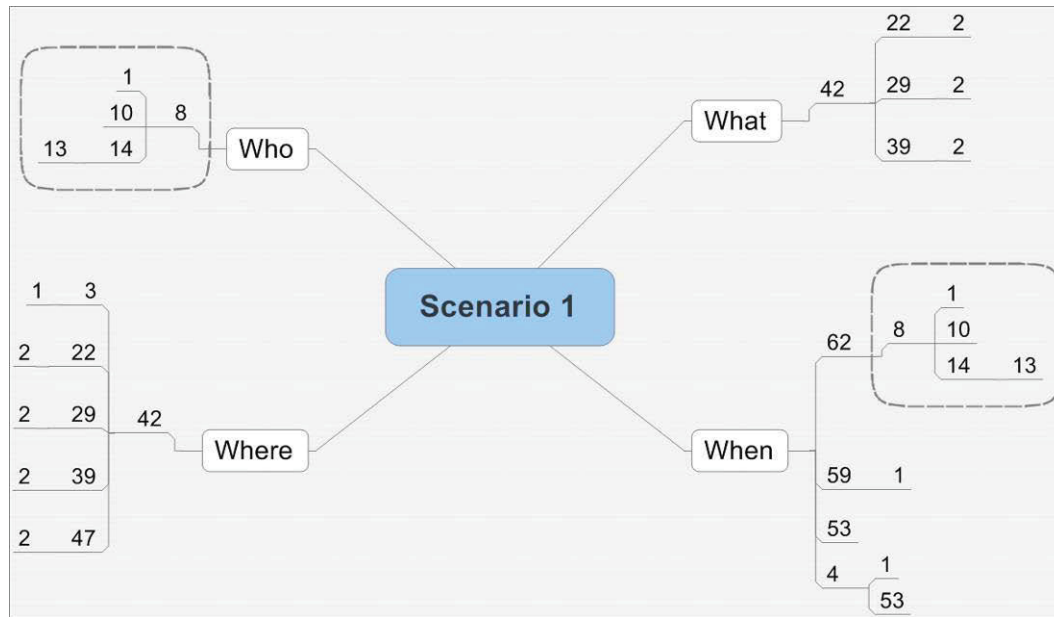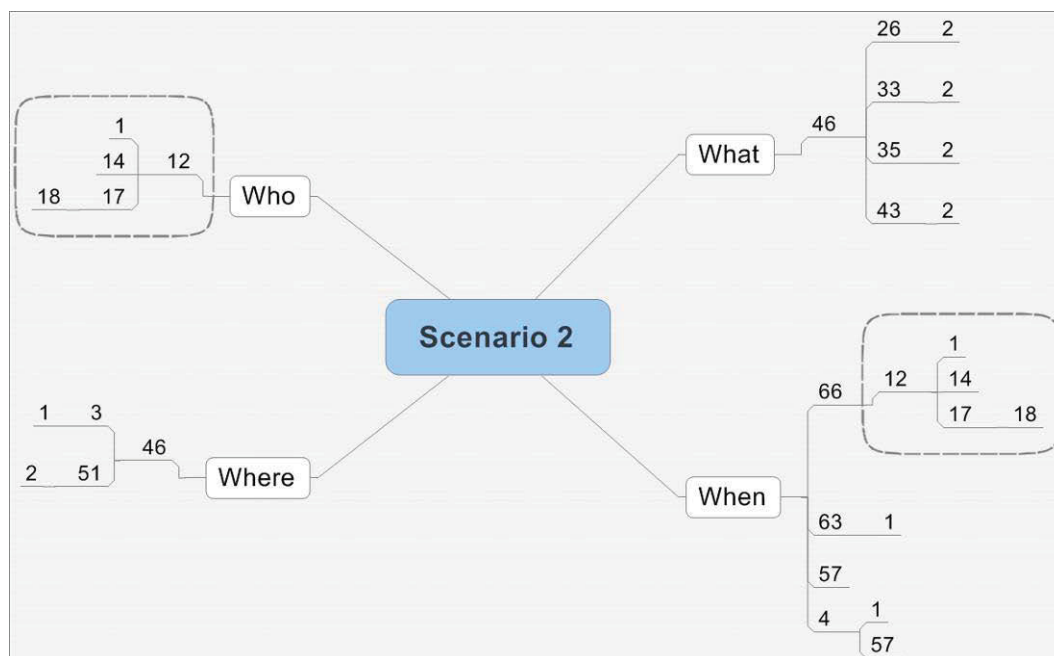
Figure 10: Scenario 1 solution tree
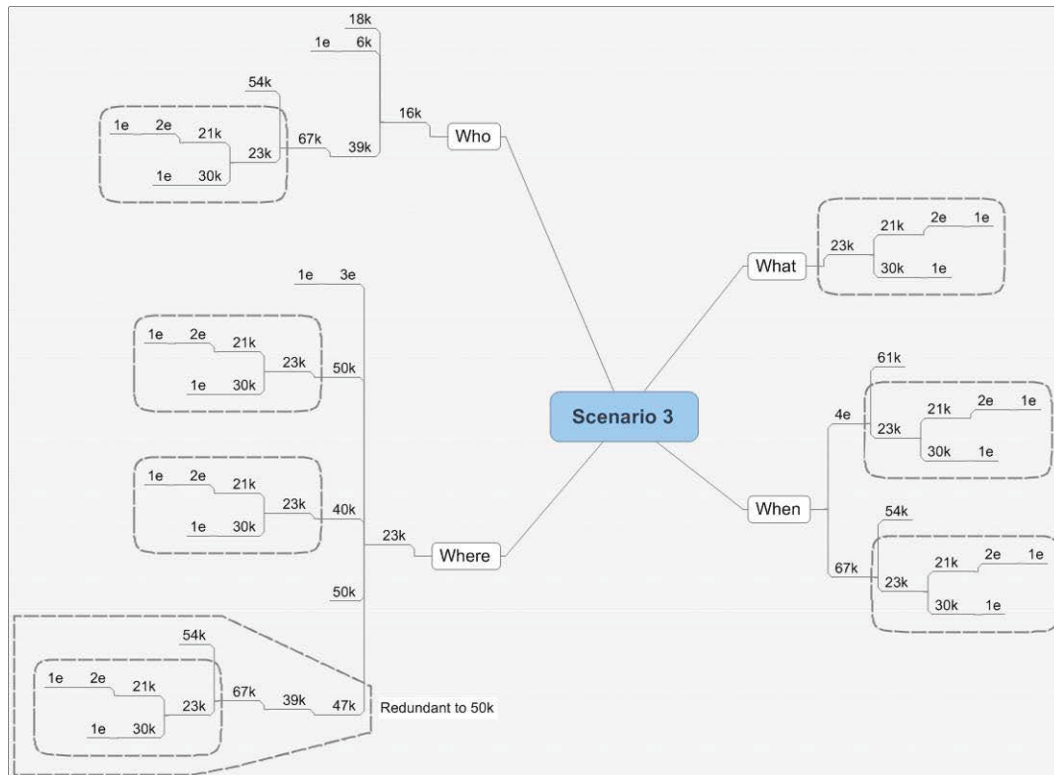

Figure 11: Scenario 2 solution tree

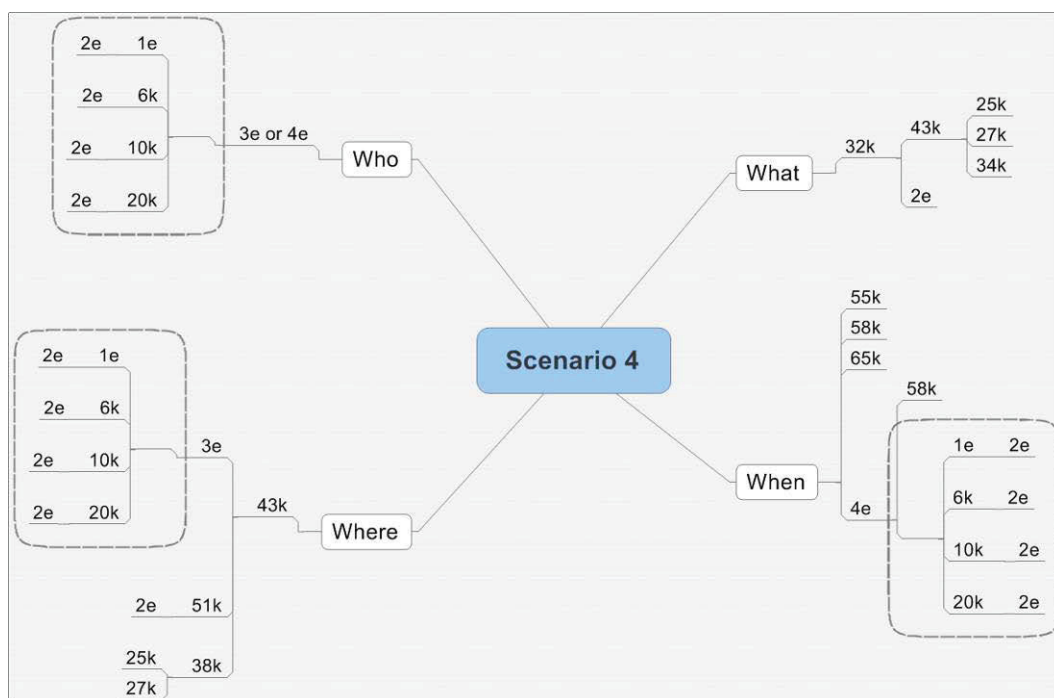**Figure 12: Scenario 3 solution tree**

Figure 13: Scenario 4 solution tree

2.13 Factoid Presentation

A limited analysis of factoid presentation effects (based on the ELICIT factoid coding) was conducted. Using the factoid keys from Scenario 1, presentation was analysed for type, subject, impact, wave, and count label. Factoids were found to be balanced for type by subject, such that each participant would receive one factoid labelled as relating to each of the four questions (Who, What, Where, When).[2]

Factoids were found to be mostly balanced for impact by subject, with each participant receiving one expert or key factoid, one supporting factoid, and two noise factoids. In Scenario 1 there is an exception in that Participant 14 receives an additional N factoid in place of an S factoid, while Participant 16 receives an additional S factoid and one fewer N factoids. This may have simply been a typo in assignment of factoids to participants, as another typo in impact classification was found for factoid 12 of Scenario 1.

Factoids were found to be balanced for impact by wave, such that as best possible factoids necessary in the completion of the solution were spread across the three waves - see Table 7. Note that the four expert factoids were all provided in the first wave.

Table 7: Factoid impact by wave

Wave	Expert	Key	Supporting	Noise	Total
1	4	5	9	16	34
2		4	4	9	17
3		4	4	9	17
Total	4	13	17	34	68

Factoids were unbalanced for type by wave in that factoids pertaining to each of four questions were not evenly distributed across presentation times. The count label included in the factoid keys was described in the software documentation as the count of the factoid within its type-impact category with the purpose of making each factoid key unique. While this description of the count label assignment appears accurate, the ordering of factoids within each counting type-impact category and the purpose of the count key is still in question. Factoid keys are rendered unique by the inclusion of the factoid number and no further explanation or reasons for the inclusion of the count label could be found.

2.14 New Scenarios

As a result of the analyses conducted to this point, the project staff had accumulated an implicit understanding of what constituted an easy versus a difficult scenario. To capitalize on this knowledge, a series of new scenarios were developed to exemplify easy and difficult scenarios.

In generating new scenarios, a number of constraints were implemented. The same distribution of E&K, S, and N factoids was used (17, 17, 34 respectively) as in the existing scenarios. E&K factoids were developed first and in such a way that by using all 17 factoids the solution could be definitively deduced. S factoids were then added to give probabilistic support to the solution. Finally N factoids were developed last to distract and detract from the solution. Note that the factoid numbers (1 through 17) used in the solution trees below reflect the E&K factoids only; once S and N factoids were developed all factoids were renumbered so as to not cue participants to the relevance of the factoid by

---

[2] Note that prior analyses had found that the type classification is not entirely accurate.

its number. The number of alternatives in each scenario was held constant with 4 Who, 4 What, 5 Where, and 4 parts to When. This scenario size is consistent with the existing ELICIT scenarios. The number of sub-alternatives was varied across question and scenario to manipulate difficulty, again consistent with the existing scenarios.

Three new scenarios were developed. The first, code named "Coffee", was designed to be as simple and easy as possible in all regards. The second, code named "Tea", was designed to be as difficult as possible through interactions across the four questions. The third, code named "Hot Chocolate", was designed to be as difficult as possible through the use of long sequential connections of logical strings.

Following scenario development, scenarios were piloted on naïve participants. The first pilot test used only the E&K factoids while the second test used all factoids. The first pilot test participant was also familiarized to the ELICIT scenarios with the existing Scenario 1, E&K factoids only. The first pilot test had one participant while the second had two participants.

For full factoid listing of the new scenarios, see Annex B.

2.14.1 Coffee

As noted, Scenario Coffee was designed to be easy as possible in all aspects. E&K factoids were designed to interact following the typical pattern seen in previous solution trees, with minimal branching and length of branches - see Figure 14. E&K factoids were as simple as possible, without explicitly stating the answer. Supporting factoids were added to lend as much support as possible in probabilistically exciting the answer, probabilistically inhibiting non-answers, and linking/delinking. Supporting factoids were limited to not independently allow deduction of the correct solution or make E&K factoids redundant. Noise factoids were designed to create as little distraction as possible and be easily dismissed as irrelevant by a participant.

The first pilot test participant solved the Coffee scenario in approximately 3 minutes, with no confusion. The second pilot test participants required approximately 10 and 15 minutes. Participants indicated that resolving this scenario was trivial with more definitive factoids and less ambiguity.
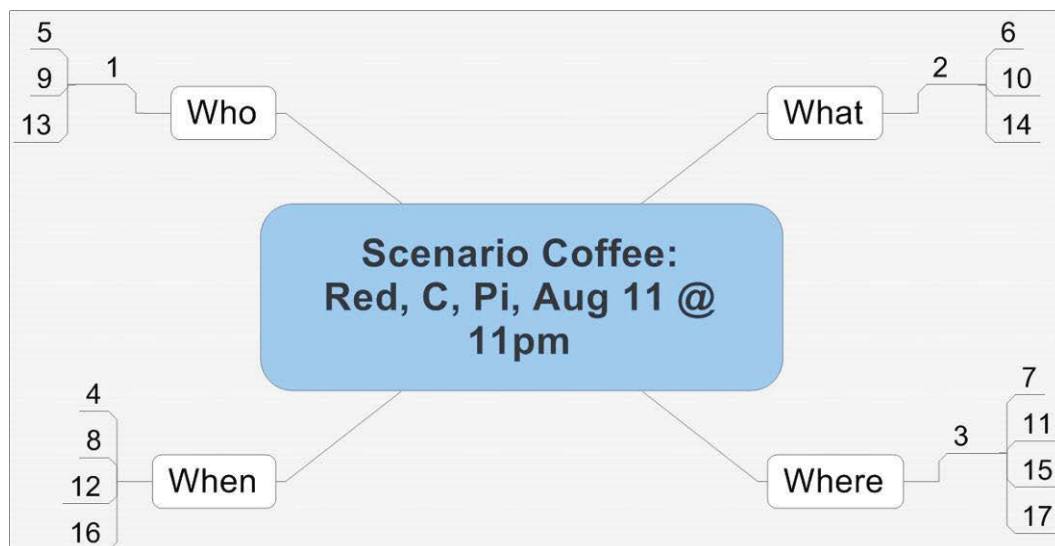

Figure 14: Scenario Coffee solution tree

2.14.2 Tea

Scenario Tea was designed to be as difficult as possible through the interaction of the Who, What, Where, and When questions. These interactions are illustrated in the solution tree as frequent re-use of interim conclusions and other factoids throughout the tree - see Figure 15; note alternative labels have been added to the solution tree for clarity. Deduction of the solution via E&K factoids was made as difficult as possible through the interdependence of questions, such that no single question could be resolved with simultaneously resolving other questions. A further complexity was introduced in that the Where answer was in fact two alternatives, with the What answer being a border crossing between these two countries. Supporting factoids were developed to offer as little support as possible while still contributing to the solution. S factoids were often non-exclusive and conditional. Noise factoids were designed to distract and detract from the solution as much as possible. N factoids probabilistically inhibited the answer, probabilistically excited non-answers, introduced new noise alternatives, and linked/delinked in meaningful ways.

The first pilot test participant required approximately 50 minutes to solve the Tea scenario; furthermore, despite experience with interaction of questions in the training scenario, the pilot participant had difficulty resolving the complexity of interaction in scenario Tea. The second test participants required approximately 35 and 60 minutes to answer this scenario and made errors in the solutions proposed. The errors were in the What and Where questions, with one participant selecting a noise alternative for What and both only identifying one of the two necessary Where alternatives. Debriefing one of the second pilot test participants revealed that a noise alternative was interpreted as satisfying one of the sub-alternatives, specifically due to past impressions about how terrorists operate. This reinforces the need to consider all possible interpretations and be wary of any potential ambiguity when constructing a scenario. Participants indicated this scenario was more difficult as factoids were less definitive and they felt more inference was required (however, in reality, this scenario can be solved without the use of inference).
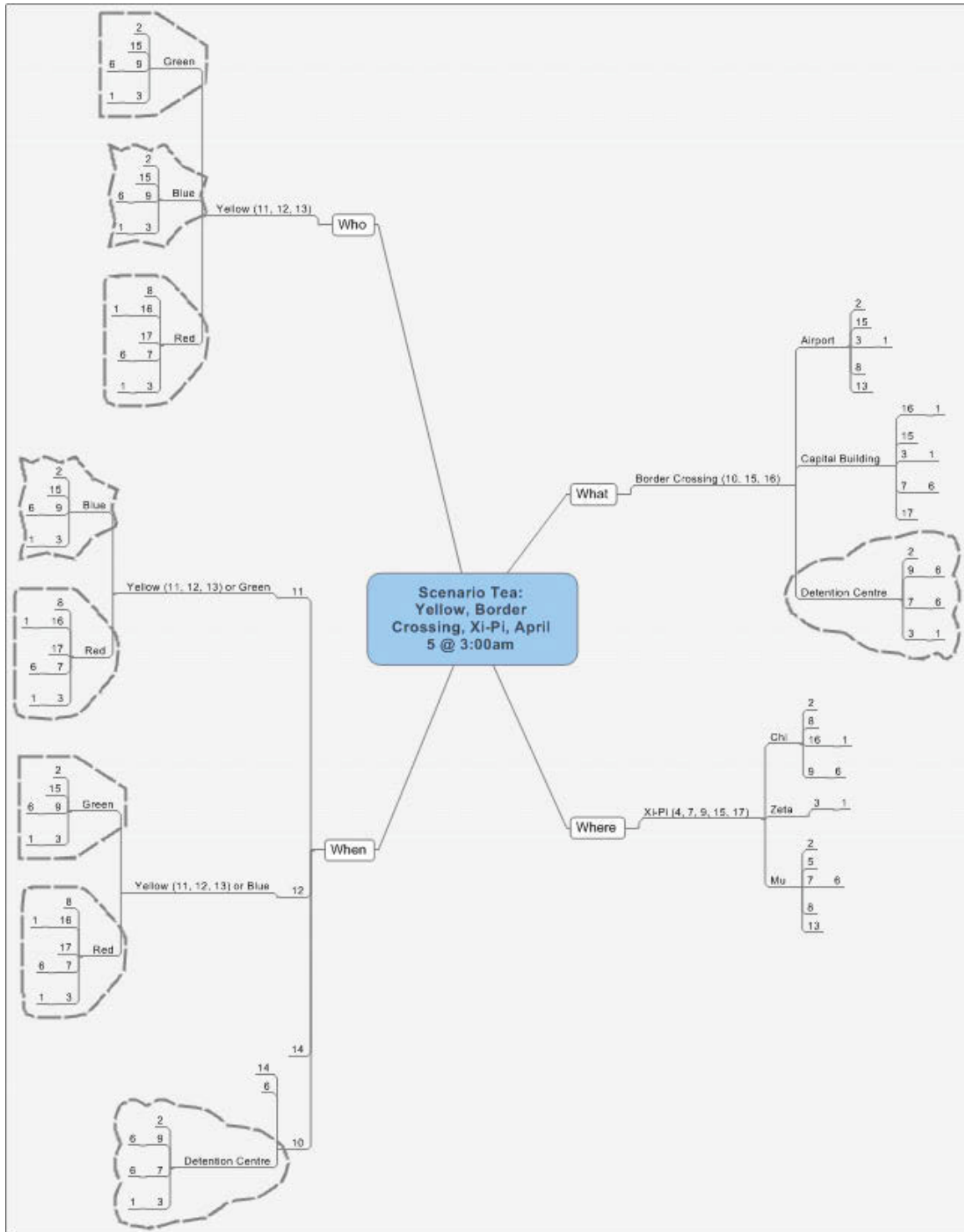
Figure 15: Scenario Tea solution tree

2.14.3 Hot Chocolate

Following the mapping of the Scenario Tea solution tree, it became apparent that its source of difficulty was the interaction across questions, as depicted by branching and the re-use of factoids and interim conclusions. A third new scenario was then developed to emphasize long sequences of conditional factoids, represented by the length of branches - see Figure 16. While many patterns of factoids are re-used throughout the solution tree, the length of branches is much longer than the previous scenarios. Again S factoids were added to provide as little support as possible while N factoids were developed to distract and detract as much as possible.

Pilot testing with the first participant revealed that no progress toward the solution could be made in approximately 20 minutes, despite utilizing the strategies successful in solving scenarios Coffee and Tea. The second pilot test did not attempt to solve the Hot Chocolate scenario.

Figure 16: Scenario Hot Chocolate solution tree

2.14.4 New Scenario Difficulty

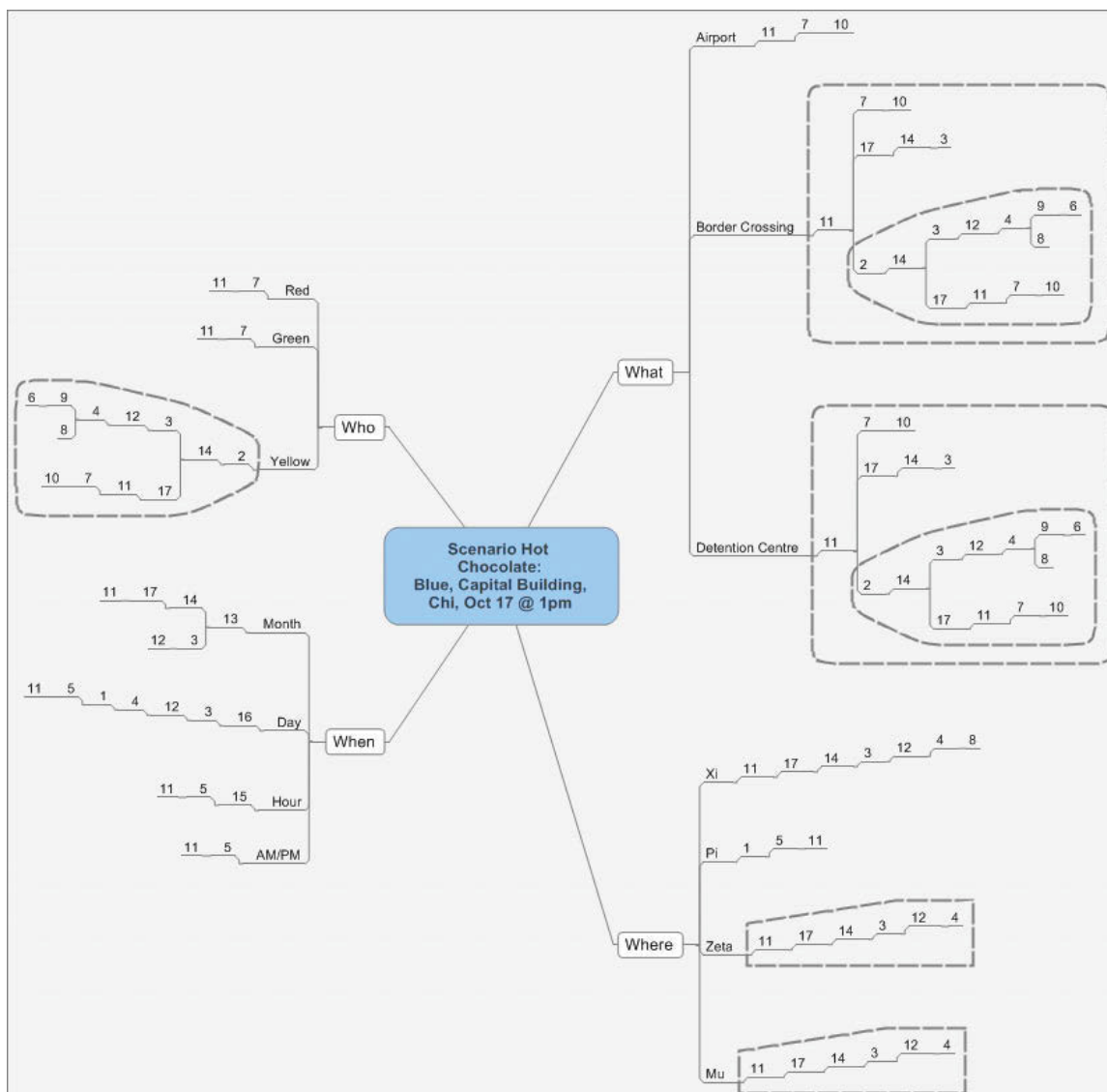Participants of the pilot tests agreed that Coffee scenario was simple and easy, while Tea scenario (and Hot Chocolate scenario) was complicated and difficult. The new scenarios were successful in creating divergent levels of difficulty; however the relative separation of the three scenarios is still an empirical question. Before efforts are taken to calibrate the relative difficulty levels, it was suggested that a moderate difficulty scenario be developed.

2.15 Four Components of Difficulty Theory

After developing and piloting the new scenarios, an effort was made to formalize the rules and principles followed in attempting to manipulate the difficulty level of the scenarios (e.g. making Coffee easy and scenarios Tea and Hot Chocolate difficult). It is proposed that scenario difficulty can be parsed into four components corresponding to the E&K factoids, S factoids, N factoids, and presentation effects. Instead of attempting to score or quantify this complexity, descriptive statements of what principles can be followed to manipulate complexity were distilled. Principles use directional phrasing where each manipulation will have a corresponding effect on scenario difficulty and doing the opposite will have the reverse effect on scenario difficulty.

2.15.1 Complexity of E&K Factoids

The first component of difficulty is the complexity of the E&K factoid interactions while deducing the correct solution. The E&K factoids are central to the total difficulty of the scenario as these factoids are the mechanism for resolution using deductive reasoning. From this analysis it was determined that E&K factoid interaction complexity can be manipulated in four areas.

First, E&K factoid interaction is more complex when within the factoid structure the information bits have greater conditionality / inter-dependency on other alternatives / sub-alternatives within question being addressed and on alternatives / sub-alternatives of other questions.

Second, E&K factoid interaction is more complex when factoids are worded as non-exclusive, in that multiple alternatives can satisfy the factoid.

Third, E&K factoid interaction is more complex when there are minimal definitive excitation or inhibition factoids, rather than more linking/delinking between factoids.

Fourth, E&K factoid interaction is more complex with size-based manipulations including the number of alternatives per question and number of sub-alternatives.

2.15.2 Strength of S Factoids

The second component of difficulty is the strength of S factoids in exciting the answer and inhibiting non-answers. As the S factoids provide stronger support, the scenario will be less difficult. The analysis of strength of S factoids also revealed four areas to manipulate.

First, S factoids provide stronger support with definitive statements than with probabilistic. Therefore as more S factoids are stated in definitive terms, the scenario will trend toward less difficult.

Second, as with E&K factoids, the exclusivity of S factoids will influence the strength of support provided by the S factoids. For excitation statements, singular exclusive statements provide stronger support to scenario resolution than non-exclusive statements; however, for inhibition statements, multiple non-exclusive statements provide stronger support.

Third, S factoids are generally stronger when dealing with alternatives than when referencing sub-alternatives. Sub-alternatives will always require linking to alternatives to be useful in determining the solution.

Fourth, excitation of the answer will provide stronger support for the solution than inhibition of non-answers, simply because there will be more non-answer than answer alternatives and the solution is phrased in answer alternatives.

2.15.3 Distraction and Detraction of N Factoids

The distraction and detraction value of N factoids is the third component of scenario difficulty. N factoids can distract from the solution providing irrelevant or unnecessary information and detraction from the solution by exciting non-answers and inhibiting answers. As N factoids become more distracting and detracting, the scenario becomes more difficult. Five areas of manipulation are proposed for N factoids.

First, N factoids are more distracting/detracting when the answer alternatives, sub-alternatives, and concepts are inhibited. Direct inhibition of the answer components will never be definitive; however, probabilistic inhibition will still disrupt inferential problem solving.

Second, excitation of non-answer alternatives, sub-alternatives, and concepts by N factoids will cause greater distraction/detraction.

Third, the introduction of new 'noise' alternatives, sub-alternatives, and concepts in the N factoids will cause distraction/detraction by enlarging the problem space.

Fourth, ambiguity in wording, specifically when different interpretations have differing implications, will increase the distraction/detraction value of the N factoid.

Fifth, the total coherence of the N factoids will influence the distraction/detraction value. Stronger coherence among N factoids will lead to more distraction/detraction than a group of incoherent disjoined N factoids.

2.15.4 Presentation Effects

Although not manipulated in the new scenarios developed, presentation effects are the final component of scenario difficulty. Four areas of presentation effects manipulation are proposed; however, the specifics of presentation manipulations are currently not as well understood.

Factoids which are conditional or dependent on other factoids or interim conclusions will be of limited use before other factoids or interim conclusions are available. In this way, it is theorized that reserving presentation of definitive factoids which are not conditional until the last presentation wave will increase the difficulty of the scenario. Withholding necessary factoids will delay the possibility of any interim or final conclusions.

Order of presentation of factoids may also influence scenario difficulty because of primacy effects. For example, the relative timing of N factoids relative to S and E&K factoids may cause incorrect hypotheses to be formed early in the scenario solving process.

The potential for a single participant (amongst all 17 participants) to use the factoids provided to them in combination to arrive at an interim conclusion will also influence scenario difficulty. The current presentation scheme allows for each participant to receive one E or K factoid and one S factoid.

Should these factoids happen to interact in a meaningful way, a participant may be able to independently arrive at an interim conclusion, thereby suggesting lower overall scenario difficulty.

While the current scenarios distribute factoids that are balanced by type (question), manipulating the types of factoids presented to individual participants may influence scenario difficulty. A participant may be better able to make judgements of relevance, create links between factoids, make inferences, and arrive at interim conclusions if the factoids provided all relate to a single question as opposed to more than one question.

Finally, in the approach used by ELICIT designers the current distribution of factoids by impact is balanced across participants. An unbalanced distribution may better enable some participants who are provided more E&K or S factoids to determine the correct solution, while others (depending on the coherence N factoids) may infer incorrect solutions.

2.16 Word Count

A participant during scenario testing suggested that the total information provided in the more difficult scenarios was higher than in the easy scenario, as evidenced by the disparity in the number of words used in respective factoids. A quick analysis of the number of words was conducted for the four scenarios provided and the three new scenarios - see Table 8 for mean (standard deviation) word counts by scenario and factoid impact. This general observation of more words in the factoids of more difficult scenarios appears to hold true for the new scenarios; however, this would suggest that all four of the existing scenarios are more difficult the Tea scenario which is unlikely based on the solution trees and four components of difficulty theory.

Table 8: Mean factoid word counts

Scenario	E Factoids (n=4)	K Factoids (n=13)	S Factoids (n=17)	N Factoids (n=34)	Overall Mean (n=28)
1	8.50 (7.05)	12.38 (3.99)	12.41 (3.54)	9.32 (3.17)	10.63 (3.94)
2	8.50 (7.33)	10.92 (3.55)	12.47 (3.32)	9.03 (3.13)	10.22 (3.79)
3	8.50 (5.20)	12.15 (4.86)	11.41 (3.47)	9.56 (3.12)	10.46 (3.81)
4	9.00 (1.15)	10.69 (2.84)	11.29 (3.16)	9.35 (2.83)	10.07 (2.93)
Coffee	12.00 (1.83)	7.15 (1.14)	9.35 (2.94)	7.35 (2.06)	8.09 (2.51)
Tea	12.00 (4.08)	9.77 (4.15)	10.47 (3.12)	8.24 (2.69)	9.31 (3.33)
Hot Chocolate	14.25 (4.57)	13.23 (4.71)	10.41 (2.53)	10.56 (3.22)	11.25 (3.64)
Total	10.39 (4.90)	10.90 (4.12)	11.12 (3.26)	9.06 (3.03)	10.00 (3.57)

It is interesting to note the consistency in mean word counts exhibited by the existing scenarios within factoid impact categories, as opposed to the much larger differences in mean word counts seen in the new scenarios within impact categories. Furthermore, there are some interesting patterns of factoid size between factoids of different impact classifications. E factoids are the shortest, followed by N factoids with K and S factoids being considerably larger for all existing scenarios. In contrast, in the new scenarios, E factoids consistently showed higher word counts, and other types of factoids among the new scenarios also showed more variability (relative to the old scenarios). These findings suggest that factoid word count is not a good indicator of scenario difficulty, but may be useful in ensuring consistent volume of text and thereby informational content.

This page intentionally left blank.

3.   Progression of Analyses

As this project has progressed, our analyses have shifted in focus and have moved to different levels of analysis. Specifically, they have moved from the level of information bits, to factoids, to whole scenarios. Within each of these levels, multiple analyses were used, but at least one approach at each level proved to be the most helpful in terms of understanding the content and/or structure of the ELICIT scenarios. The analyses conducted at each level are shown in Figure 17, with the most helpful analysis within each level highlighted within the blue arrow at the right.
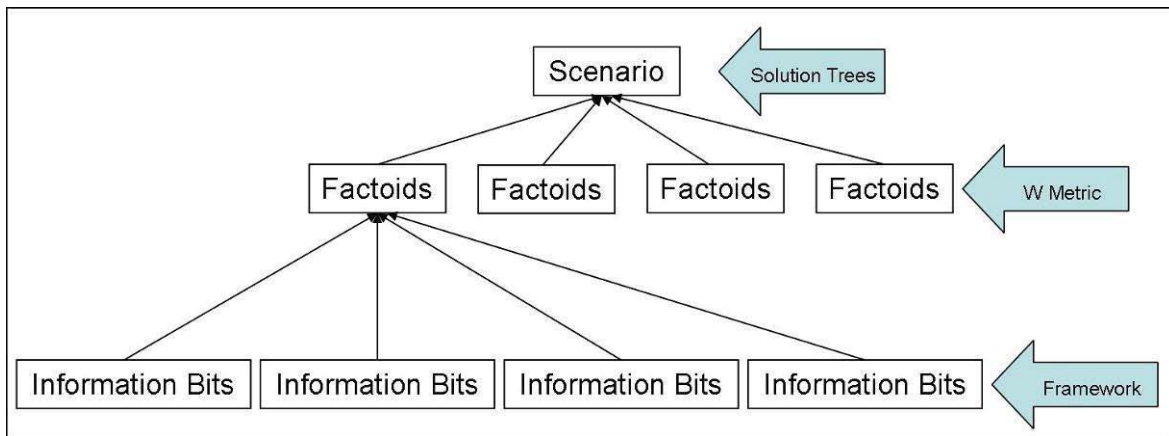

Figure 17: Progression of analyses

At the level of discrete bits of information, the framework scoring allows for the quantification of the absolute value of information bits from all types of factoids. The information bit scoring enables factoid valuation within the scenario; however, no meaningful differences were found across scenarios.

At the level of factoid, the W metric appears to be the most sophisticated measure of E&K factoid contributions to the solution, as per the solving matrices. This metric is currently limited to E&K factoids and also does not enable comparisons across scenarios.

At the level of scenarios as a whole, the solution trees present a visual representation of E&K factoid interaction in determination of the solution. While no quantifiable metrics have been developed for the solution trees as of yet, the salient visual features facilitate comparison across scenarios. Also of note are the four components that influence scenario difficulty. In our view, being able to control these components systematically would enable the generation of scenarios of varying difficulty for future research.

This page intentionally left blank.

4.   Further Analyses

This project has explored the nature of the existing ELICIT scenarios in order to begin to dimensionalize them as a precursor to being able to systematically control them. This control is critical to being able to quantify the difficulty of a given scenario for future research and experimentation. To this end, the analyses conducted hopefully represent a good start and offer a way forward for future research effort.

The current project has progressed the understanding of scenario difficulty and strategies for calibration, however the analysis work in the development of a team scenario content generation framework is by no means complete. Based on the knowledge gained during the completion of this project, discussions with the project team and SA, several ideas for future analyses that might prove fruitful are offered.

A scoring system for supporting factoids (similar to the W metric using the matrix approach) may be useful in evaluating the strength and distribution of support provided. By making the assumption that the probabilistic statements of the S factoids are true, strength and distribution of support could be measured using similar techniques as the W metric.

Another observation made during the analyses was the noise factoids in the current ELICIT scenarios vary widely in their content and quality as representatives of the "noise" category. A scoring system for noise factoids could also use the matrix approach to evaluate the coherence of the noise. The rationale is that more coherent noise that presents a more viable alternative will be of greater distraction and detraction to the participants. By developing scoring systems for S and N factoids, scoring of factoids by impact could be extended to a total scenario scoring and thereby allow cross scenario comparisons.

One potential area of improvements is in the concept of noise. Noise in psychological signal detection theory refers to anything that is not a true signal or target. In this way noise should refer to the number of non-answer alternatives. The current N factoids are in fact a mixture of "filler" factoids which contain information of zero value in determining the solution and "undermining" factoids which are of negative value. Formally defining these two types of factoids, previously known as noise, will allow a new potential manipulation of scenario difficulty by the ratio of "filler" to "undermining" factoids.

Another pathway to pursue in developing a whole scenario scoring systems would be the scoring of solution trees. By quantifying the branching and length characteristics of the solution trees, mathematical comparisons of the difficulty, complexity, and size of scenarios could be possible.

New scenarios could be developed to further populate the difficulty continuum. Specifically, two new scenarios are proposed. The first would take the same approach as used in the Hot Chocolate scenario and creating logical strings but with the Coffee scenario end goal of creating the easiest scenario possible. The second would seek to find the middle difficulty, half way between the new scenarios at either extreme.

A scoring system for the four components of difficulty could be developed based on the current understanding of difficulty. This scoring system, as well as others previously developed, could then be validated with behavioural measures such as time to completion and information exchange.

As new scenarios are developed and the nature (and quantification) of difficulty are better understood, there may be opportunity to better refine the scenario structure without changing the facilitation software. As this work proceeds, it will also be important to test any new scenarios and to validate the assumption underlying the scenario (e.g. that a scenario created to have a high level of difficulty is actually more difficult to solve). This will require a range of participants and multiple forms of feedback.

5. References

Parity Communications Inc. (2008). ELICIT - The Experimental Laboratory for Investigating Collaboration, Information-sharing and Trust, ELICIT 2.2 Web-based Software Guide.

This page intentionally left blank.

Annex A - Analysis Files

Factoid analysis of set 1.xls

Factoid analysis of set 2.xls

Factoid analysis of set 3.xls

Factoid analysis of set 4.xls

Annex B - New Scenario Files

Scenario Coffee.xls

Scenario Tea.xls

Scenario Hot Chocolate.xls

DOCUMENT CONTROL DATA		
(Security classification of the title, body of abstract and indexing annotation must be entered when the overall document is classified)		
1. ORIGINATOR (The name and address of the organization preparing the document, Organizations for whom the document was prepared, e.g. Centre sponsoring a contractor's document, or tasking agency, are entered in section 8.) Publishing: DRDC Toronto Performing: Humansystems® Incorporated, 111 Farquhar Street, Guelph, ON N1H 3N4 , Monitoring: Contracting:		2. SECURITY CLASSIFICATION (Overall security classification of the document including special warning terms if applicable.) UNCLASSIFIED
3. TITLE (The complete document title as indicated on the title page. Its classification is indicated by the appropriate abbreviation (S, C, R, or U) in parenthesis at the end of the title) Development of a Team Scenario Content Generation Framework (U) Élaboration d'un cadre de génération de contenu pour les scénarios d'équipe (U)		
4. AUTHORS (First name, middle initial and last name. If military, show rank, e.g. Maj. John E. Doe.) Andrew Morton; Barbara D. Adams		
5. DATE OF PUBLICATION (Month and year of publication of document.) April 2011	6a NO. OF PAGES (Total containing information, including Annexes, Appendices, etc.) 56	6b. NO. OF REFS (Total cited in document.) 1
7. DESCRIPTIVE NOTES (The category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of document, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.) Contract Report		
8. SPONSORING ACTIVITY (The names of the department project office or laboratory sponsoring the research and development – include address.) Sponsoring: Tasking:		
9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant under which the document was written. Please specify whether project or grant.)		9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.) W7711–088128/001/TOR
10a. ORIGINATOR'S DOCUMENT NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document) DRDC Toronto CR 2010–105		10b. OTHER DOCUMENT NO(s). (Any other numbers under which may be assigned this document either by the originator or by the sponsor.) 8128–11
11. DOCUMENT AVAILABILITY (Any limitations on the dissemination of the document, other than those imposed by security classification.) Unlimited distribution		
12. DOCUMENT ANNOUNCEMENT (Any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11), However, when further distribution (beyond the audience specified in (11) is possible, a wider announcement audience may be selected.)) Unlimited announcement		

DOCUMENT CONTROL DATA (Security classification of the title, body of abstract and indexing annotation must be entered when the overall document is classified)

13. ABSTRACT (A brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual.)

(U) Web–based software called ELICIT (Experimental Laboratory for the Investigation of Collaboration, Information–sharing, and Trust; Parity Communications Inc., 2008) offers a research platform for exploring team performance dynamics, and command and control (C2) organizations. This project aimed to use existing ELICIT scenarios as a starting point for developing a formal relational/structural framework that would enable the creation of a means of systematically generating the content required for complex team scenarios. This knowledge could be instructive for designing a workable approach to use for tightly controlled future scenario generation.

ELICIT scenarios contain 68 short statements, typically a single sentence, called factoids. Analyses conducted in this course of this project slowly progressed from decomposition of the ELICIT factoids into bits of information, to the level of factoids as a whole, to the level of the scenario as a whole. A number of key observations were made through this process.

The primary observation is that while the scenarios have an underlying structure which allows both deductive reasoning and inference, there seems to be little ability for the systematic creation of factoids or manipulation of scenario difficulty. Through the analyses conducted there appears to be a basic scenario structure upon which a framework for the understanding of difficulty, complexity, and other potentially interesting experimental manipulations could be built. The analyses of this project represent a start toward the goal of better understanding (and hence controlling) scenario characteristics and systematic scenario creation.

(U) Le logiciel web portant le nom d'ELICIT (Experimental Laboratory for the Investigation of Collaboration, Information–sharing, and Trust, de Parity Communications Inc., 2008) offre une plateforme de recherche pour l'exploration de la dynamique de la performance collective et des organisations de commandement et contrôle (C2). Ce projet visait à utiliser les scénarios ELICIT existants comme point de départ pour l'élaboration d'un cadre relationnel/structurel formel qui permettrait la création d'un moyen de créer systématiquement le contenu nécessaire pour les scénarios d'équipe de type complexe. Cette connaissance pourrait être instructive pour concevoir une approche apte à être mise en pratique pouvant être utilisée pour la production future de scénarios strictement contrôlés.

Les scénarios ELICIT contiennent 68 courts énoncés, habituellement formulés en une seule phrase, appelés « factoïdes ». Des analyses menées au cours de ce projet ont progressées lentement, de la décomposition des factoïdes d'ELICIT en segments d'information jusqu'au niveau des factoïdes dans leur ensemble, puis au niveau du scénario dans son ensemble. Un certain nombre d'observations clés ont été faites au cours de ce processus.

La principale observation touche le fait que bien que les scénarios possèdent une structure sous–jacente qui permet de faire à la fois des raisonnements déductifs et de l'inférence, la création systématique de factoïdes et la manipulation du niveau de difficulté des scénarios semblent difficiles. Les analyses effectuées ont permis de dégager ce qui semble être une structure de scénario de base à partir de laquelle il serait possible de bâtir un cadre permettant de comprendre la difficulté et la complexité des scénarios ainsi que d'effectuer d'autres manipulations expérimentales potentiellement intéressantes. Les analyses de ce projet constituent un point de départ vers une meilleure compréhension (et ainsi un meilleur contrôle) des caractéristiques d'un scénario et de la création

systématique de scénarios.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (Technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus, e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

(U) Development; team; scenario; framework

UNCLASSIFIED

Defence R&D Canada

Canada's Leader in Defence
and National Security
Science and Technology

R & D pour la défense Canada

Chef de file au Canada en matière
de science et de technologie pour
la défense et la sécurité nationale

DEFENCE **R&D** DÉFENSE

www.drdc-rddc.gc.ca