

This paper was prepared by Anthony Judge of the UIA. The on-line version can be viewed at <http://www.laetusinpraesens.org/docs00s/questdb.php>. The UIA's "Questions" database can be viewed at <http://www.diversitas.org/db/x.php?dbcode=qu>. See also the illustrations in documents 10.1.2. and 10.1.4. in this volume.

Background

The experiment described below follows from an initial interest of the German Research Centre for Artificial Intelligence (DFKI), in support of the questions project of the international nonprofit organization Dropping Knowledge – as clarified during a workshop on the online databases of the Union of International Associations (Saarbrücken, 8 December 2005). Dropping Knowledge subsequently appropriated this information as the basis for establishing an online web facility to enable people worldwide to ask questions and to be exposed to answers – thereby creating a "Living Library". The categorization of the questions was undertaken using the ontology developed by the UIA (cf *Enabling a Living Library*, 2006). (See also "Complementary Knowledge Analysis / Mapping Process" in this volume.)

The concern here, in contrast, was whether it was possible to generate a Questions database from three long-established UIA databases: World Problems-Issues, Global Strategies-Solutions, and Human Values. These databases are part of the online *Encyclopedia of World Problems and Human Potential*, originally initiated in collaboration with Mankind 2000, whose development was most recently funded by the European Commission. The databases are integrated with others on international organizations, international meetings, biographies and bibliographies (cf *Yearbook of International Organizations, International Congress Calendar*).

The thousands of problems, strategies and values identified from the documented preoccupations of the network of international organizations (governmental and nongovernmental) provide a relatively objective focus for the generation of questions associated with those preoccupations – or implicit in them. Clearly a particular interest in this experiment is to determine in what ways the result of generating questions could be meaningful and significant. The work builds on the possibilities of the use of such databases for simulations (cf *Simulating a Global Brain: using networks of international organizations, world problems, strategies, and values*, 2001).

WH-questions

There is an extensive literature on what are termed "WH-questions". "WH-questions" refer to questions of the type: How? Why? Where? What. Which? When? Who? Further comments on studies in relation to such questions are noted below.

The Questions database was first generated experimentally in December 2005, and then more comprehensively in October 2006. In each case by applying a template of the WH-questions to the titles of Problems, Strategies and Values.. This can be done by embedding the "seed title" (XXXX) in a suitable template phrase. For example:

- *How is XXXX caused?*
- *Who is responsible for YYYY?*
- *Where does XXX occur?*
- *When does XXXX occur?*
- *What is XXXX?*

The range of templates is illustrated by Table 1.

As is clear from Table 1, different templates were used both according to the source database and according to the WH-Question. Since many Problems and Strategies have a number of alternative titles (notably employing synonyms), these too have been used as seeds for the generation of alternative titles for a question – effectively constituting alternative formulations of the same question (but clustered together in the Question entry). Although they may be accessed through their keywords, they are not treated as distinctly profiled questions.

Preliminary results

The very preliminary results in generating these questions in December 2005 are indicated in Tables 2, 3 and 4. The subsequent generation of the Questions in October 2006 gave the results in Table 5.

Remarks

The results given here are of course extremely preliminary. Some clarifications regarding the tables are appropriate:

- The application of a single WH-template of each of the 7 types (Who, Where, etc) to each seed entity gives rise to the column labelled "main".
- The application of any additional WH-templates (6 in the case of Problems) to each seed entity gives rise to the column labelled "WH-variants".
- The results for "main" and "WH-variants" may be the same (as in the case of Strategies) because the number of templates applicable is the same.
- The "alternative titles" do not give rise to separate question entries for statistical purposes.
- The labelling of the functional relationships is in practice different in the case of Problems and Strategies although in each case the systemic notion of facilitating or constraining is the same.

The results could be substantively affected by:

- increasing or reducing the number of original entities selected for application of the WH-templates;
- increasing or reducing the number of WH-question templates themselves;
- culling questions after generation.

Table 1: Templates used experimentally to generate questions

Source database	Generated query (**= not finally used)			
	WH-query	Phrase-1	Seed	Phrase2
World Problems – Issues (13 templates applied)	How	much	XXXX	is there?
	How	does	XXXX	happen? (**)
	How	is	XXXX	caused? (**)
	Why	does	XXXX	happen?
	Why	give priority to	XXXX	?
	Why	does God allow	XXXX	to happen? (**)
	Why	be concerned by	XXXX	? (**)
	Where	does	XXXX	occur?
	What	is	XXXX	?
	What	causes	XXXX	?
	What	results in	XXXX	? (**)
	Who	causes	XXXX	?
	Who	is responsible for	XXXX	?
	Who	is concerned about	XXXX	?
	When	does	XXXX	occur?
	When	will	XXXX	occur?
	When	did	XXXX	arise?
	Which	kind of	XXXX	?
Global Strategies – Solutions (14 templates applied)	How	can	XXXX	be enabled?
	Why	is	XXXX	unsuccessful?
	Why	give priority to	XXXX	?
	Where	is	XXXX	undertaken?
	Where	is	XXXX	successful?
	What	is required for	XXXX	?
	What	causes	XXXX	to fail?
	Who	undertakes	XXXX	?
	Who	is responsible for	XXXX	?
	Who	is concerned about	XXXX	?
	When	is	XXXX	undertaken?
	When	will be	XXXX	undertaken?
	When	was	XXXX	undertaken?
	Which	kind of	XXXX	?
Human Values – constructive or destructive (9 templates applied)	How	is	XXXX	elicited?
	Why	is	XXXX	valued?
	Why	give priority to	XXXX	?
	Where	is	XXXX	found?
	What	is	XXXX	?
	Who	exemplifies	XXXX	?
	Who	values	XXXX	?
	When	is	XXXX	evident?
	Which	kind of	XXXX	?
Human Values – polarities (10 templates applied)	How	are	XXXX	related?
	How	can	XXXX	be reconciled?
	How	can	XXXX	be transcended?
	Why	is the	XXXX	relation so
	Where	are	XXXX	reconciled?
	What	transcends	XXXX	?
	Who	embodies	XXXX	?
	Who	exemplifies the	XXXX	ambiguity?
	When	are	XXXX	transcended?
	Which	kind of	XXXX	relationship?

Table 2: Preliminary results (December 2005)					
	Seed entities		WH-templates used	Questions generated	
	Total	Selected		Main	WH-Variants
Problems	59205	12995	13	168935	239252
Strategies	42032	12848	14	179872	167426
Values	3257	3209	9 /10	29111	16470
Totals	104494	29052		377918	423148
					801066

Table 3: Indication of distribution of seed entities by type												
	Problems-Issues						Strategies-Solutions					
	Profiles			Links			Profiles			Links		
	1996	2000	%	1996	2000	%	1996	2000	%	1996	2000	%
A	0	196	n.a.	0	3,507	n.a.	0	1,518	n.a.	0	16,767	n.a.
B	170	187	10%	5,300	7,090	34%	158	154	-3%	3,697	4,253	15%
C	575	722	26%	13,816	19,347	40%	1,100	1,089	-1%	17,096	25,206	47%
D	2,162	2,740	27%	30,613	52,451	71%	3,315	3,452	4%	19,374	43,329	124%
E	3,857	5,378	39%	29,626	52,587	78%	3,008	5,298	76%	11,092	50,677	357%
F	3,072	3,917	28%	38,625	61,604	59%	1,382	1,972	43%	7,015	19,580	179%
G	2,153	30,279	1306%	5,979	47,112	688%	7,685	13,107	71%	3,604	69,059	1,816%
Other	214	12,716	5,842%	905	26,255	2,801%	12,850	6,105	-52%	61,129	34,070	-44%
Total	12,203	56,135	360%	124,864	269,953	116%	29,498	32,695	11%	123,007	262,941	114%

Table 4: Indication of seed entity relationships							
	Hierarchical links			Functional links			
	<i>Broader</i>	<i>Narrower</i>	<i>Related</i>	<i>Aggravating</i>	<i>Aggravated by</i>	<i>Reducing</i>	<i>Reduced by</i>
Problems	26403	35500	14264	31024	31105	1507	1529
Strategies	27134	32541	3010	3302	2902	17826	16911
Values		11392					

Table 5: Final results of question generation (October 2005)

	Problems		Strategies		Values		Value-Polarities		Totals		
	Main	WH-Variants	Main	WH-Variants	Main	WH-Variants	Main	WH-Variants	Main	WH-Variants	All
Seed entities	45892		31055		2978		229		80154		
WH-templates	7	6	7	7	7	2	7	3	28	18	46
Main	321244	-	217385	-	20846	-	1603	-	561078	-	561078
(Alternative titles)	133917	114786	107422	107422	9807	2802	-	-	-	-	-
WH-Variants	-	275352	-	217385	-	5956	-	687	-	499380	499380
Total questions	321244	275352	217385	217385	20846	5956	1603	687	561078	499380	1060458
Broader	433244	371352	370174	370174	79765	22790	0	687	883183	765003	1648186
Narrower	346577	297066	284396	284396	0	0	79744	34176	710717	615638	1326355
Related	113302	97116	33768	33768	0	0	0	34863	147070	165747	312817
Total hierarchical	893123	765534	688338	688338	79765	22790	79744	69726	1740970	926888	2667858
Aggravates	235347	201726	42196	42196	77	22	0	0	277620	243944	521564
Aggravated by	233989	200562	42637	42637	0	0	0	0	276626	243199	519825
Reduces	11298	9684	174125	174125	0	0	0	0	185423	183809	369232
Reduced by	11340	9720	175469	175469	0	0	0	0	186809	185189	371998
Total functional	491974	421692	434427	434427	77	22	0	0	926478	856141	1782619
Strategies	163289	139962	0	0	336966	96276	182	78	500437	236316	736753
Problems	0	0	163205	163205	256585	73310	20090	8610	439880	245125	685005
Values	223293	191394	337134	337134	0	0	0	0	560427	528528	1088955
Total cross-database	386582	331356	500339	500339	593551	169586	20272	8688	1500744	1009969	2510713
Total relationships	1771679	1518582	1623104	1623104	673393	192398	100016	78414	4168192	3412498	7580690

Table 6: Format of a displayed question record

Output / Displayed Record	Source database: The questions have been generated from the titles of entries in three different databases: World Problems-Issues (P), Global Strategies-Solutions (S), or Human Values (V)
	Seed title: Questions have been generated by taking each of the (possibly several alternative) titles of a Problem, a Strategy or a Value. The title associated with this entry is indicated here
	Type code: In the source database (Problems, Strategies, Values), each entry is allocated a type code. Typically, in the case of Problems or Strategies, the lowest letters of the alphabet indicate the most generic entries where the higher letters in the alphabet indicate more specific entries. In the case of the Values database, entries of Type C are associated with Constructive Values, those of Type D with Destructive Values and those of Type P with Value Polarities
	WH-Question type: The following classical types of generic "WH-questions" are used to generate the Questions in this database: 1=When? 2=Where? 3=Which? 4=How? 5=What? 6=Who? or 7=Why? The relevant one for this Question entry is indicated here.
	Question variant: To distinguish between the (possibly several) alternative titles, each is given a single digit number. The first (1) is that which is presented as the principal title of the entry in the source database. Questions of different types (What? Where? etc) are applied to each such title which maintains that single digit number.
	WH-Question family: These are the titles of all the Questions generated from a single title of the Problem, Strategy or Value entry from which the Question derived. They therefore all share the same "Question variant" (1-9) but are all associated with a different WH-Question type (Who? Where? When? Which? How? Who? Why?)
Relationships	Broader questions: These are Questions that are more general, or more contextual, than that of the entry. They correspond to Questions that have been generated, as appropriate, from any broader Problem, Strategy or Value in the corresponding databases.
	Narrower questions: These are Questions that are more specific than that of the entry. They correspond to Questions that have been generated, as appropriate, from any narrower Problem, Strategy or Value in the corresponding databases.
	Related questions: These are Questions that are associated in some non-specific way with the Question of this entry. They correspond to related entities in the seed entry in the source database whether a Problem or a Strategy.
	Aggravates (P) / Constrains (S): These are Questions that may impose constraints on those of the same type with which they are linked in this way. They correspond to Questions that have been generated, as appropriate, from any equivalent Problem or Strategy in the corresponding databases.
	Aggravated by (P) / Constrained by (S): These are Questions that may be constrained in some way by those of the same type with which they are linked in this way. They correspond to Questions that have been generated, as appropriate, from any equivalent Problem or Strategy in the corresponding databases.
	Reduces (P) / Facilitates (S): These are Questions that may reduce or facilitate those of the same type with which they are linked in this way. They correspond to Questions that have been generated, as appropriate, from any equivalent Problem or Strategy in the corresponding databases.
	Reduced by (P) / Facilitated by (S): These are Questions that may be facilitated in some way by those of the same type with which they are linked in this way. They correspond to Questions that have been generated, as appropriate, from any equivalent Problem or Strategy in the corresponding databases.
	Strategy questions: These are Questions generated from any Strategy entry associated with the Problem entry from which this Question was generated in the Problems source database. This field is not relevant in the case of Questions generated from the Strategies database.
	Problem questions: These are Questions generated from any Strategies entry associated with the Strategy entry from which this Question was generated in the Strategies source database. This field is not relevant in the case of Questions generated from the Problems database.
	Value questions: These are Questions generated from any Values entry associated with the Problem or Strategy entry from which this Question was generated. It is not relevant in the case of entries from the Values database.

Integration into UIA set of databases

The Questions database, with its 1,060,458 Questions, was integrated into the UIA set of online databases by Tomáš J. Fülöpp in October 2006. It is freely accessible over the web. This integration has the advantage of using the common search and visualization interfaces developed for the other databases (including World Problems, Global Strategies, Human Values, International Organizations, International Meetings, etc). The format of a displayed question record is given in Table 6. Information on various types of relationships between questions clearly depends on the presence of such information in the seed entry in the source database.

Comments

Fundamental to this exploration are the following issues:

- Problems as themselves constituting a form of "question" calling for some form of "answer" – although, as "questions", they may themselves merit questioning, thus clarifying a challenging problem by asking more searching questions
- Strategies as themselves constituting a form of "answer" – although, as "answers" that may themselves merit questioning – by "putting them to the question"
- Values motivating both acknowledgement of any "question" or sensitivity to the need for an "answer" – but whose underlying concern may also be fruitfully questioned

The original datasets of Problems, Strategies and Values – as developed since 1972 – have large numbers of relationships between records within and between those databases. These relationships have been the subject of extensive "hyperlink editing", most recently by Nadia McLaren, enabling extensive analysis (cf Anthony Judge and Nadia McLaren, *Feedback Loop Analysis in the Encyclopedia Project*, Extract from the final report on *Information Context for Biodiversity Conservation*, 2000). From the above, given that this pattern has been preserved, it can be seen that these are based on various types of relationship:

- Hierarchical
- Systemic or functional
- Cross-database

The generated Questions database is therefore an overlay of the above network of links in terms of:

- The 7 WH-questions
- The variants (max. 3) of some of those WH-questions

The generated questions are linked back to the entity from which they were generated, whether a problem, a strategy or a value – thus providing another point of access to these datasets. This integration of the other databases through specially framed questions may prove to be particularly valuable.

The question templates are necessarily different between databases and between variants. The simplistic nature of the templates may not necessarily result in question titles that are grammatically totally correct at this stage – but sufficiently so for the purpose of this experiment.

Possibilities from this approach

This work was inspired at an early stage by that of Stafford Beer, Syd Howell, Alan Mossman, and Gordon Pask who developed a set of techniques on the occasion of a conference on *Improving the Human Condition: Quality and Stability in Social Systems* (Silver Anniversary International Meeting, London, 1979) of the Society for General Systems Research (SGSR). The resulting documents, tables and maps were presented in *Metaconferencing possibilities: Discovering people / viewpoint networks in conferences* (1980). Of particular relevance is their early use of a question-statement refinement technique and mapping of the results as applied to an international conference involving people well-disposed towards such techniques.

A particular exploration touching on WH-questions was made in various interrelated papers prior to this experiment:

- Functional Complementarity of Higher Order Questions: psycho-social sustainability modelled by coordinated movement (2005)
- Engaging with Questions of Higher Order: cognitive vigilance required for higher degrees of twistedness (2005)
- Question Avoidance, Evasion, Aversion and Phobia: why we are unable to escape from traps (2006)
- Conformality of 7 WH-questions to 7 Elementary Catastrophes (2006)
- Cognitive Feel for Cognitive Catastrophes: Question Conformality (2006)

Some interesting work could be done to refine the WH-question templates and to explore their functional relationships (as suggested in the above papers).

The hierarchical linkages between questions would provide a very interesting technique for moving to more generic questions or "drilling" down to more specific questions.

The functional links offer an interesting possibility for exploring learning pathways based on questions. In this context the detection and exploration of loops of questions (using network analysis techniques) raise many points of interest (cf the work of Ron Atkin on simplicial complexes).

Additional features might include:

- A facility for users to impose their own preferred template on the seed entities
- It is possible that it would be relatively easy to apply automatic translation techniques to such questions to obtain linguistic variants – raising issues of the correspondence to WH-questions in other languages and the possibility of other forms of question not envisaged in this experiment.

Once the formatting is clarified, the issue is whether the database in its entirety lends itself to interesting analysis, notably with the visualization and other tools (already offered as options to the online presentation of search results). Some thoughts are:

- Since we are dealing with questions, does the network of linkages make sense as a learning pathway anyway (Broader, Narrower, etc)?

-
- Can the questions be clustered in interesting ways, given that both the Problems and Strategies were analyzed to discover interesting loops?
 - Are the links between questions from different databases meaningful?
 - What would it take to move beyond, or filter out, those questions that could be identified as more contrived or artificial?
 - How does this kind of experiment help to focus attention on "better" questions?
 - To what extent does the pattern of questions (with the templates used) usefully covers the forms of interrogation with regard to XXXX?
 - Given the artificial templates used to generate the questions, is any meaning mainly to be derived at more abstract levels of analysis?
 - Given the recurring significance to contemporary challenges of governance of "problems", "strategies" and "values" – from which the "questions" were generated – is there further significance to be obtained from any explicit links to the "organizations" and "meetings" where their understanding and importance are clarified? Are "questions" then to be usefully considered as the underlying focus of "organizations", but especially of "meetings"?

Is there a case for recognizing that it is not new "answers" to old "problems" which is the fundamental challenge, but rather new "questions" with respect to those old "problems"?

Question of significance?

The key question with any visualization of a pattern of information, such as that enabled by this experiment, is whether it offers any new insight. Clearly unusual patterns of information can be generated, but do they lead to unusual insights?

The visualization metaphors used have the advantage of holding disparate questions in relationship to one another in a manner which suggests the possibility of a more integrative perspective. Interacting with the various visualization options may enhance that possibility. At issue are the conditions under which such a perspective might emerge as more than an artifice of the design metaphor. More fundamental is the issue of what forms of visualization enable new insight and how they may compare with those explored with respect to these patterns of relationship between questions.

Of particular interest is the extent to which the patterns of questions may be compared with those characteristic of stages in learning pathways and individuation processes (cf George Siemens, *Connectivism: Learning as Network-Creation*, 2005).

