



12 December 2005

## Preliminary Attempt at Generating Questions from UIA Databases: Problems, Strategies, Values

-- proof of concept exercise --

Nadia McLaren, Tomáš J. Fülöpp and Anthony Judge  
[nadia@uia.be](mailto:nadia@uia.be) / [tomi@uia.be](mailto:tomi@uia.be) / [judge@uia.be](mailto:judge@uia.be)

### Summary

Approximately 800,000 questions were experimentally generated from almost 30,000 entities (Problems, Strategies and Values). It is tentatively estimated that approximately 5.5 million links (or learning pathways) could be generated between these questions in the light of the existing hierarchical and functional links between the entities in the databases of the [Union of International Associations](#) (UIA).

### Background

This is a response to the expressed interest of the [German Research Centre for Artificial Intelligence](#) (DFKI) in the databases of the [Encyclopedia of World Problems and Human Potential](#), notably with respect to the project of [dropping knowledge](#). This was clarified during a workshop in Saarbrücken (8 December 2005). [dropping knowledge](#) envisages providing an online web facility to enable people worldwide to share knowledge, ask questions and to be exposed to answers, of which a test version is already operational.

The following assessment is about how the entity titles in the UIA Problems, Strategies and Values databases could be used to generate questions as the possible basis of a new UIA Questions database.

### WH-questions

There is an extensive literature on what are termed “WH-questions”. These are questions of the type: How? Why? Where? What? Which? When? Who? They are fundamental to a number of aspects of web search engine technology and to the particular interests and expertise of DFKI.

An interesting example of use of these questions is the query technique offered by [googlism.com](#) and the answers it offers in response to user input.



Exploration touching specifically on these questions was recently made by Anthony Judge (UIA) in two related papers:

- [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)

In response to the [dropping knowledge](#) project, the set of WH-questions could be applied to each Problem, Strategy or Value title in the UIA databases. This can be done, for example, by embedding the title in a suitable template phrase:

- What causes XXXX?
- Who is responsible for XXXX?
- When does XXXX occur?
- Why give priority to XXXX?
- Where is YYYY undertaken?
- What causes YYYY to fail?
- Why is ZZZZ valued?
- What is ZZZZ?
- Etc

## Preliminary results

A summary of the contents and development of the Problems and Strategies databases is shown in the table below (at the end of an EU INFO2000 contract 1997-2000). This is the raw material from which the questions were generated on 12 December 2005 using the WH-question templates and selected types of entity from each database.

Statistical overview of selected Encyclopedia databases												
Types	Problems – Issues						Strategies—Solutions					
	Profiles			Links			Profiles			Links		
	1996	2000	% growth	1996	2000	% growth	1996	2000	% growth	1996	2000	% growth
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%



Very preliminary results in generating these questions are given in the following table. "Primary" questions were derived from main names of entities; "Variants" from alternate names of entities.

Overview of questions generated					
Database (types)	Entities		WH- Question templates	Questions generated	
	Total	Selected	Applied	Primary	Variants
<b>Problems (B-F)</b>	59,205	12,995	13	168,935	239,252
<b>Strategies (B-F)</b>	42,032	12,848	14	179,872	167,426
<b>Values (C, D, P)</b>	3,257	3,209	9 /10	29,111	16,470
<b>Totals</b>	104,494	29,052		377,918	423,148
				<b>Grand total</b>	<b>801,066</b>

The next table illustrates the potential domain of linkages that could be generated between generated questions, based on existing links within the Problems, Strategies and Values databases.

Number of original entity links available for conversion into links between questions							
	Hierarchical links			Functional links			
	Broader	Narrower	Related	Aggravating / Facilitating	Aggravated by / Facilitated by	Reducing / Alleviating	Reduced by / Alleviated by
<b>Problems</b>	26363	35508	14261	31024	31105	1507	1529
<b>Strategies</b>	27134	32541	3010	3302	2902	17826	16911
<b>Values</b>	11619	11392					
<b>Totals</b>	65116	79441	17271	34326	34007	19333	18440
						<b>Grand total</b>	<b>267934</b>

With respect to the possible links between questions, a more detailed set of estimates is provided, by type, in an appended table.



## Remarks

The above results are of course extremely preliminary and are designed simply as a proof of concept exercise.

The results could be substantively affected and improved by:

- Increasing or reducing the number of original entities (a selection of section types different from those indicated) according to careful criteria on level of desired detail. Indications of the numbers of omitted entities of other types, and the consequences for generation of links, are given in the appended detailed table
- Increasing or reducing the number and/or style of WH-question templates, and possibly refining the manner in which they are applied
- Whereas in this exercise "variant" question names were not treated as unique questions, it is possible that they would make separate but related questions, instead of alternate expressions of the same question.
- Culling questions after generation
- Grammatical improvements (notably adjustment for singular and plural when embedding entity names in a template) requiring more sophisticated algorithms or "eyeballing".

The above results do not (yet) reflect the potential generation of *links between* the generated questions on the basis of the original links between entities in the individual databases. The exercise focused on the ability to generate the questions, not on issues arising from generation of the links.

## Possibilities from this approach

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 generated questions would of course be linked back to the entity from which they were generated, whether a problem, a strategy or a value.

The hierarchical linkages between questions would provide a very interesting remote learning technique for moving to more generic questions or down to more specific questions, or vice versa: in effect along the information—knowledge—wisdom axis.

It is clear that by excluding or including types of entity the user can be exposed to simpler (or more fundamental) questions (at the top of hierarchies) or exposed to more detailed questions (if that is the user's point of entry).

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). The detection of (strategically) "central" questions is of particular interest. It is



perhaps appropriate to note that recognized learning pathways, between Modes of Awareness within particular traditions, feature as functional links between entities in the UIA's Human Development database (not considered in this preliminary Question database exercise).

The links from the Questions database to the Problems and Strategies databases—themselves linked to entities in the Organizations database—could be used to "confront" Organizations with the questions with which they are seemingly concerned. This process might be used to elicit questions from them, or to amend the questions proposed to them.

Additional links between Questions could be derived from "handmade links" *between* the Problems, Strategies and Values.

It is possible that it would be relatively easy to apply automatic translation techniques, possibly supplemented by human translation, to such questions to obtain linguistic variants. This could generate interest in WH-question templates that would be of greater relevance in other languages.

Given the way in which Question entities are generated, it is also possible that they could be generated in real time in response to user interest—rather than constituting a separate database.

The [dropping knowledge](#) agenda is to solicit / elicit questions. How could such harvested questions be fruitfully confronted by this generated database?

Given the repository of linked materials constituted by the UIA databases, how could the possible generation of questions be used to "jump start" and "populate" the [dropping knowledge](#) system?

Just as [dropping knowledge](#) specifically intends to link text materials with appropriate images, it is worth noting that some UIA profiles of entities display images, such as logos and geographic maps in the case of Organizations, and topological maps in the case of all databases.

### **Afterthought: Relevance of investigations at a general systems conference**

Discussions in [Metaconferencing possibilities: Discovering people / viewpoint networks in conferences \(1980\)](#) are based on documents, tables and maps prepared by cyberneticians Stafford Beer, Syd Howell, Alan Mossman, and Gordon Pask on the occasion of a conference on *Improving the Human Condition: Quality and Stability in Social Systems* -- the Silver Anniversary International Meeting (London, 1979) of the Society for General Systems Research (SGSR).

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.



**Preliminary estimates of number of "Links between Questions" generatable from UIA Problems, Strategies and Values  
for proof of concept particular types of entities were selected / omitted on 12-12-2005**

	Types	Hierarchical relations			Functional relations				Links				
		Broader	Narrower	Related	Totals	Aggravates / Facilitates	Aggravated by / Facilitated by	Reduces	Reduced by	Totals	Original entity links	WH templates	Possible question links
<b>Strategies</b>	selected	171	252	68	2822	48	60	332	179	619	6263	14	87682
		1115	8976	287	11963	599	394	3932	2362	7287	31213	14	436982
		3655	9446	798	18950	931	819	5664	5232	12646	50546	14	707644
		5673	8298	1339	21294	1107	1053	5687	6500	14347	56935	14	797090
		2234	3819	518	7656	617	576	2211	2638	6042	21354	14	298956
		<b>12848</b>	<b>32541</b>	<b>3010</b>	<b>62685</b>	<b>3302</b>	<b>2902</b>	<b>17826</b>	<b>16911</b>	<b>40941</b>	<b>166311</b>	<b>70</b>	<b>2328354</b>
<b>Problems</b>	omitted	1487	1095	288	17142	84	18	314	105	521	34805	14	487270
		13773	4576	1427	26795	1747	1984	5105	6268	15104	68694	14	961716
		13924	4970	390	8903	980	1286	1949	1892	6107	23913	14	334782
		<b>29184</b>	<b>26857</b>	<b>2105</b>	<b>52840</b>	<b>2811</b>	<b>3288</b>	<b>7368</b>	<b>8265</b>	<b>21732</b>	<b>127412</b>	<b>42</b>	<b>1783768</b>
		187	302	595	3444	847	1290	33	59	2229	9117	13	118521
		732	1457	1119	7794	3037	3358	129	157	6681	22269	13	289497
<b>Values</b>	selected	2741	6186	3338	19819	8174	8603	338	416	17531	57169	13	743197
		5445	10892	4324	24193	7685	7636	344	320	15985	64371	13	836823
		3939	7526	4885	20882	11281	10218	663	577	22739	64503	13	838539
		<b>13044</b>	<b>26363</b>	<b>14261</b>	<b>76132</b>	<b>31024</b>	<b>31105</b>	<b>1507</b>	<b>1529</b>	<b>65165</b>	<b>217429</b>	<b>65</b>	<b>2826577</b>
	omitted	196	152	2749	3247	44	33	1	2	80	6574	13	85462
		30434	5725	1504	40236	1942	1745	67	60	3814	84286	13	1095718
<b>Values</b>	selected	15531	13171	445	22178	709	795	43	34	1581	45937	13	597181
		<b>46161</b>	<b>46330</b>	<b>4698</b>	<b>65661</b>	<b>2695</b>	<b>2573</b>	<b>111</b>	<b>96</b>	<b>5475</b>	<b>136797</b>	<b>39</b>	<b>1778361</b>
		225	11392	0	11617	0	0	0	0	0	23234	10	232340
		987	4148	0	4148	0	0	0	0	0	8296	9	74664
		1992	7246	0	7246	0	0	0	0	0	14492	9	130428
		<b>3204</b>	<b>11619</b>	<b>0</b>	<b>23011</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>46022</b>	<b>28</b>	<b>437432</b>
	<b>45</b>	<b>0</b>	<b>225</b>	<b>225</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>450</b>	<b>0</b>	<b>0</b>	
<b>Total selected</b>		<b>29096</b>	<b>79441</b>	<b>17271</b>	<b>161828</b>	<b>34326</b>	<b>34007</b>	<b>19333</b>	<b>18440</b>	<b>106106</b>	<b>429762</b>	<b>163</b>	<b>5923363</b>
<b>Total all</b>		<b>104486</b>	<b>118177</b>	<b>24074</b>	<b>280554</b>	<b>39832</b>	<b>39868</b>	<b>26812</b>	<b>26801</b>	<b>133313</b>	<b>694421</b>	<b>244</b>	<b>9154492</b>