WORKING NOTE AC103

Alan Creak
30 September 1996

ANALYSING WORDKEYS

These are some notes on a seminar given by Marianne Hickey at York[1]. My intention was to restate Marianne's approach in terms of my analysis, in the hope of illuminating both. And, while I was at it, I added some comments.

Marianne suggested that "the interface between a non-speaking person and a communication aid can be viewed in two parts – a physical and a language interface" (quoted from the abstract of her seminar), and described WordKeys, a communication aid in which prestored sentences are retrieved by single-word keys using content search without explicit encoding.

I have used my informal notation[2] to describe the sequence of actions which might be performed in a WordKeys system to produce one sentence. While doing so, I tried using a slightly extended version of the notation, refined to deal with this example. The extended version is better than the original, so at least I've gained something from the exercise. This is the result :

Item name	*Locus*	*Vocabulary*	*Description*	*Remark*	*Interface*
S-notion	S	Thoughts	What S (the sender) wants to say.		
→ keyword	S	Any word(s)	Chosen by S as a selection key.	An approximation. How easy is it ?	
→ typing-plan	S	Keyboard actions	How to type the chosen word.		P ?
→ typing	S - W	Keyboard actions	The S – W (WordKeys) interface.		P
→ keyword	W	Word			L
→ keyword set	W	Word list	Related words, homophones, etc.	Automatic expansion	L
→ sentences	W	Sentence list	Contain instances of the keyword set.	Selected from a stored collection.	L
→ ranked	W	Sentence list	The sentence list sorted.	In what order ?	L
→ display	W - S	Dialogue box	Select from window – scrolling, maybe.	Defines sentence vocabulary.	P
→ ranked	S	Sentence list	S has read the sentences.	Uses sentence vocabulary.	
→ S~sentence	S	Sentence	S chooses a sentence for communication.	An approximation.	
→ selection-plan	S	Dialogue selection	How to select the chosen sentence.		P ?
→ selection	S - W	Dialogue selection		Number ? Clicking ?	P
→ S~sentence	W	Sentence			
→ S~sentence	W - H	Language	H is the hearer.		
→ S~sentence	H	Language	Received by H.		
→ message understood	H	Thoughts	Understood by H.	Not necessarily S-notion.	

The central notion of the table is that any communication can be described as a sequences of transformations of the originator's material through a sequence of vocabularies and, eventually, through a sequence of interacting entities, which might be people, interface components, wires, computer programmes, and so on. The sequence is described in the first column of the table, while the place in the

system and the vocabulary are presented in the next two. The names used are arbitrary. The remaining columns are essentially commentary; I remark further on the final column later.

The transformation sequence is a model of the communication process, and one would expect it to work better the more accurate the description. In practice, I've found it helpful even if I'm frankly guessing the mechanism; the discipline of having to invent some plausible sequence of events at least clarifies some aspects of the process, and makes one conscious of things that must happen somewhere even if the details remain obscure.

THE INTERFACE.

This is only a partial description. I haven't carried the decomposition of the actions to the lowest possible level (planning individual key operations, etc.), because I don't think that anything is gained by doing so in this example. I haven't included anything about possible iteration – in practice, for example, after looking at the initial selection of sentences presented by WordKeys, S might choose to use a different keyword or to add an additional keyword in order to change or refine the selection. That would add to the complexity, but Marianne said that it wasn't much used. Neither have I included anything about building up the collection of sentences or specifying additional keywords, but a complete description should take such supporting activities into account.

In the rightmost column of the table I've recorded my guesses at how each step should be seen as part of the physical (P) or language (L) interfaces. It makes sense, insofar as the two notations are consistent, and the unclassified bits are clearly associated with one or other of the communicating entities rather than with their interaction. That said, I'm not sure that the interface parts alone are very informative. Perhaps that's because I don't know anything about the context of the interface approach.

My analysis perhaps makes it clearer that there are two separate interaction episodes. During the first, the keyword is translated into a set of sentences, while in the second a specific sentence is chosen.

I've labelled two actions "P ?". Both actions are of the same type; they are both cases of the sender planning what to do with the interface. While the actions themselves are clearly performed by the sender alone, they cannot be performed without information about the physical interface, and I'm not sure how they would fit into the two-parts-to-the-interface model. In my treatment, I can introduce an explicit interaction in which the sender acquires details of the acceptable vocabulary from the interface; this information can come from the interface itself or it can be conveyed in manuals or other documentation. In the first interaction episode of this example, the vocabulary is that of ordinary typing, which I've assumed is understood, but in the second it is a combination of a low level vocabulary for selecting an item from a dialogue box (also assumed understood) and a high level vocabulary of sentences, conveyed by the "display" step.

SELECTING A SENTENCE.

The intention of my analysis is to identify the steps in the overall communication process, so that each can be checked for effectiveness, difficulty, speed, etc.

So far as the possibility of the process is concerned, there are only two questionable steps : the selection of the keyword, and the selection of the sentence from the list presented by WordKeys. (All right, so it was fairly obvious from the start – but I find it reassuring to be able to write down a plausible guess at the overall process to convince myself that all the bits really fit together.) In each of those cases, some sort of approximation is involved, and to ensure satisfactory communication the combined effects of these approximations must be small.

It is clear from the first step that this is not going to be easy to achieve in general, for the first step is a drastic approximation : a single word, however carefully chosen, can occur in an unbounded number of sentences. It is therefore up to the second step to restore the required content in an effective way. To do so, WordKeys, given a keyword, must be able to produce in response a set of sentences which in some sense covers all the things one might reasonably want to say about the keyword. If this set of sentences is large, then it should also be possible to sort them in such a way that there is a good chance of finding the sentence required quickly – in essence, a sort on popularity.

Is that feasible ? The obvious answer is "no". The reason for this answer is that the set of things one might want to say about most words is very large, and selecting the desired sentence from the large set (assuming that it's there at all) is inevitably difficult. There are two cases (I think *only* two) in which this reason might be challenged. First, if the set for some reason is small, then the selection might be much easier. Second, if many sentences from the set are equally acceptable, then *an acceptable* sentence (rather than *the required* sentence) is likely to be much easier to find. Marianne's examples come into these classes.

In interactional encounters, it might well be that any sentence which more or less appears to be in the context of the current exchange is acceptable as the next utterance. If a casual conversation is about (say) chickens, then almost any sentence about chickens can appear to be in context. Consider "I like roast chicken"; "Eggs are expensive these days"; "Poultry farms are rather smelly"; "I like watching ducklings swim" : all are likely to be accepted in a conversation as at worst minor changes in topic. But that only applies so long as the conversation remains casual. If it develops into a more directed discussion of the evils of battery farming, random interjections will be much less appropriate.

In transactional encounters, it is commonly rather clear from the start what will have to be said, and a fairly convincing script (in the artificial intelligence sense) can be drawn up in advance. In this case, the set of possible utterances might be quite small, so once again selection is easy. Marianne used the example of arranging car insurance, where there are a smallish number of well defined points which have to be covered somewhere, and the order is not of great importance.

Unless other special cases turn up, therefore, the trick in achieving more fluent conversation on specific topics is likely to be to take advantage of one of these two special cases. One can aim either to constrain the topic as far as possible, or to relax the sentence selection criterion. It is not at all clear how the second can be achieved except by trivialising the conversation, and that strategy might not always be welcome.

Attempts to constrain the topic are therefore likely to prove more fruitful. As an example, one might provide for temporary default keywords which are always included in the search profile when defined – so to engage in a concentrated discussion on battery farming, one could enter "poultry farm" as a keyword which remains fixed for the duration. The additional keywords could be either added to the set of words required in the sentence during the process of selection from the stored collection, or used in sorting the selected sentences for presentation.

(In pondering these questions, I developed a rather fanciful geometrical analogy which gave me some insight of uncertain quality into the selection process. I decided that it was rather too far-fetched to include here, but I wrote it down anyway[6].)

CAN YOU PREPARE FOR INTERACTIONAL ENCOUNTERS ?

A more direct approach mentioned by Marianne is to provide means for the person using the system to prepare beforehand sets of sentences which could be explicitly linked to some specified keyword as well as relying on the usual search for keywords. In effect, the intention is to use the "transaction" technique of preparing potentially useful material in the context of an interactional encounter. In her example, the word "holiday" was explicitly made a keyword for the sentences :

"We went to France last summer for two weeks."
"Eight of us shared a cottage in Brittany."
"I am planning to go skiing – if I can afford it !"
"There were lots of small beaches nearby."
"We went swimming nearly every day to cool off !"

As with my "chicken" examples, each of those would be quite acceptable at any time as a remark in a desultory conversation about holidays, and one can imagine more directed conversations in which all sentences would figure very naturally –

"We went to France last summer for two weeks."
"That sounds exciting !"
"Eight of us shared a cottage in Brittany."
"Was it a nice place ?"

"There were lots of small beaches nearby."
"How was the weather ?"
"We went swimming nearly every day to cool off !"
"What will you do next year ?"
"I am planning to go skiing – if I can afford it !"

Even though the responses are not always direct answers to the other person's remarks, the conversation reads naturally. But that's an illusion, for in real life you can't programme your conversation partner to make the right responses. How about :

"We went to France last summer for two weeks."
"How did you get there ?"
or *"Do you speak French ?"*
or *"How long for ?"*

In terms of my description, WordKeys would flounder with no useful sentence in the vocabulary shown in the display. I guess, on the basis of no solid evidence whatever[6], that there are so many possible, and quite sensible, responses to the original sentence that it would be extraordinarily hard to prepare for more than a tiny fraction of them.

This isn't just a quibble. Without being well acquainted with the interactional-transactional classification of conversations, I'd guess that this one is interactional. It doesn't have a set topic, or a well defined purpose, but people naturally ask questions, and expect sensible answers. Even without direct questions, one expects to be able to stick to a topic for a few exchanges, so you can't just keep changing the subject and regard the result as a real conversation. (Leon made that point in the seminar.) I think it's an important point, because – while trivial conversations are perhaps better than none – a directed conversation is much more interesting. "Most speech encounters interactional", it says in the seminar notes[1], but if my argument is correct that's only part of the story, and the interesting conversations fall between the two "easy" classes that I defined.

Without knowing a lot more about (perhaps) discourse analysis (or maybe something else) than I do know, I can't offer any useful further comment, but the argument does suggest that there's a danger of keeping things going by trivialisation rather than by intelligent sentence selection. That wouldn't encourage potential conversation partners.

COGNITIVE AND PHYSICAL LOAD.

By inspecting the transformations listed in the first column of the table, one can make some guesses at the load imposed on the sender in selecting a sentence. (I'm working on a more formal version of the method[5], which should permit much more detailed descriptions and possibly simulations, and therefore more satisfactory estimates of the load.)

The missing step – which is the particularly clever bit of WordKeys – is the cognitive step of remembering the correct code for the sentence you wish to retrieve. As compared with many other sentence-retrieval methods, this is a big advantage. But is there any counterbalancing cognitively difficult step which could nullify the advantage ?

The candidates are the keyword selection, comprehending the sentence list, and selecting a sentence for output. There are other cognitive operations, but these are common in other systems too, so the ones I've listed are the discriminating factors. All I can say about those is that, except for the two special cases where (if I'm right) the problem is unusually simple, they all look potentially hard, and likely to get harder as the collection of sentences grows. Unless the keyword, or set of keywords, is very precisely chosen (which might itself be hard), the selected set of sentences will be large, which makes it hard to search through and decide on the best choice.

The task is not simplified by the requirement to attend to the continuing conversation in parallel with the effort of driving the interface. What is the effect of reading several other sentences on one's comprehension of a conversation ?

SPEED.

WordKeys is in competition with other methods such as semantic compaction for sentence selection, word prediction for accelerated typing, and others. I'll compare it with word prediction, which will in any case have to be provided to cope with cases where no appropriate stored sentences are available.

It's clear that when WordKeys works, it is likely to work very well, returning a whole sentence for the cost of typing one word (or enough of it to predict the rest) and a selection from the list of sentences. But what happens when it doesn't work ? This will be the case if either no sensible keyword comes to mind, or if the chosen keyword doesn't retrieve a relevant sentence.

With a small number of stored sentences, it might be possible to remember them in enough detail to decide quickly whether any is appropriate in the current context, and if so what keyword might be useful. In the first case, an immediate return to the word predictor is sensible, while in the second the keyword should work. The worst damage is the time taken to decide that no sentence will do, and even that might not be severe with a limited repertoire.

With more sentences, this decision will be more difficult to make, particularly if the collection has been accumulated over a long time so that the early sentences, and the words used, have been forgotten. In this case, if anything like a plausible keyword comes to mind, it's sensible to try it (otherwise there's not much point in using WordKeys at all); then, if it works, all is well. But if it doesn't work, a significant amount of time has been lost, and the result is slower than the word predictor alone.

Whether there's an overall gain in speed depends on statistics, and I've no idea how to work it out without further information. Experiments are indicated.

SHOULD I EXPAND MY NOTATION ?

I ask the question as a continuation of the previous paragraph, because I haven't previously wanted to describe anything quite so probabilistic. I could expand the question a little into something like "Should I be able to work it out without further information ? – and, if so, should I adapt the notation to incorporate whatever's needed to manage the working out ?".

The answer to the first part is "don't know, but with a strong bias towards no". The calculation seems to depend on values for the probability that the collection contains an acceptable sentence and the probability that it will contain a word from the expanded keyword set; if several sentences were selected, one would also have to take into account the amount of list searching required, so the ranking technique would be important. It seems to me unlikely that this sort of information can be defined with sufficient generality to be useful in a general treatment. It might also be important to take into account the possibility of misunderstanding; the sentence chosen might not mean the same to the hearer as it did to the sender, and if it was already an approximation the combined shift in meaning might be significant.

I think that the answer to the second part is "no". That's certainly the preferred answer in this case ! The method is intended as a way of describing systematically what happens in a communication so that possible points of difficulty can be identified, and it seems to do that reasonably well. The more formal notation in development[5] might make it possible to build a model of the system, but it will be a simulation model, and will require explicit data to work with; it won't give general results.

Given the required information, statistical or otherwise, it might be possible to use the qualitative model as a basis for developing methods which could give useful quantitative or general results, but I don't see that such a development would affect the usefulness of the information given by the model as it stands.

POSSIBLE EXTENSIONS TO WORDKEYS.

(These comments are suggested by remarks made during the seminar.)

There was some discussion at the seminar on identifying context from previous utterances; such methods would be included in the category of techniques used to constrain the topic of the conversation. (Marianne didn't say much about it, but it wasn't a new idea[4].) It seems likely, though, that a completely automatic method might in some circumstances be inconveniently slow to respond; some manual control

of the current effective keyword set should be retained so that any sudden change in topic could be handled simply.

As an extension to that approach, perhaps it would be even better to include also the parts of the conversation contributed by other participants. While speech recognition would be of significant help in this respect, manual entry of default keywords would also allow for this technique to be used to some extent. It seems that taking account of the whole of a conversation is of value in word prediction systems[3], so it isn't unreasonable to expect it to be useful in predicting useful sentences as well.

The question of modifying stored sentences was also raised. A single stored sentence could in principle be used to generate a large number of others by simple linguistic transformations of its number, gender, mood, tense, and so on. The stored sentence "I like roast chicken" could by this means be transformed into "They liked roast chicken", "I don't like roast chicken", "Do you like roast chicken ?", and so on. If it worked, it would multiply the available vocabulary manyfold. The example of Eliza comes to mind; while much of its apparent "intelligence" is the result of ingenious contrivance and the brilliant use of the model of the psychiatrist, Eliza demonstrates the effectiveness of such transformations – though one would hope for rather more subtlety in an adaptation for WordKeys.

It is not so clear that such a method would be universally easy to use. The cognitive demands of a set of transformations which appear to require a fairly sophisticated understanding of language might be high. Two cognitive actions are necessary : first, a promising model sentence must be identified, even though it is in an inappropriate form; and, second, the required transformations must be identified and applied. On the other hand, it might turn out that such tasks are not really as demanding as they appear, and are comparatively easy to learn; I don't know.

REFERENCES.

1 : M. Hickey : *Speech for non-speaking people – the application of communication modelling and text-based intelligent retrieval,* Seminar (with distributed notes), York Computer Science Department, 23-ix-1996.

2 : G.A. Creak : *Reaching Beyond Words In Rehabilitation Computer Systems*, unpublished Working Note AC96 (May, 1996).

3 : J.L. Arnott, J.M. Hannan, R.J. Woodburn : "Linguistic prediction for disabled users of computer-mediated communication", *ECART 2 : Proceedings of the European Conference on Rehabilitation Technology*, 26-28 May, 1993, Stockholm, Sweden, paper number 11.1.

4 : M. Hickey, C. Page : "Polyvox : flexible message selection in a communication prosthesis for non-speakers", *ECART 2 : Proceedings of the European Conference on Rehabilitation Technology,* 26-28 May, 1993, Stockholm, Sweden, paper number 11.3.

5 : G.A. Creak : *Interactors in rehabilitation system design*, unpublished Working Note AC97 (August, 1996).

6 : G.A. Creak : *Sentences that mean "X", more or less*, unpublished Working Note AC104 (September, 1996).