Testing and demo in Lüneburg

Test data base: Where you can make experiments and try them out without modifying the production data base:

Input: http://l435.dsv.su.se/eu/kom/faqprod1;login

Usage: http://l562.dsv.su.se/w4h-faq-en

Production data base: Where your changes will affect the set of FAQs for our future users:

Input: http://cmc.dsv.su.se/eu/kom/faqprod1;login

Usage: http://web4health.info/en/answers/

Classification

The basis of the natural-language question-answering system:

1. Produce, for every FAQ, one or more question templates, corresponding to questions which users might ask and where this FAQ might be suitable.

2. Encode the templates in such a way that also variations in the user queries, such as use of synonyms, will still match the template.

Example 1:

Title: Obesity and risks for health

Examples of user queries:

Is obesity dangerous?

What are the risks with obesity?

How is morbidity influenced by obesity?

Why should I go on a diet?

Encoding of example 1:

Required: $obesity $bulimia ; $bad risk risks danger dangers dangerous morbid* illness* mortal* / why $should ; $slim

Forbidden:

Optional: $health

Priority: risk risks danger dangers dangerous $obesity morbid* illness* mortal*

Top-level syntax:

Required: obesit* bulimi* ; risk risks danger dangers dangerous morbid* illness* mortal*

This will match any user query which contains one of the words before the ";" and one of the words after the ";".

You can have any number of ";", for example:

Required: treat* ; overweight* ; surgic*

Synoyms:

It is important to find all reasonable synonyms to words, which users may use. For example, the FAQ "Risks with obesity" might be classified as:

Required: obesit* ; risk risks

but then the following question will not find this answer:

"What are the risks of being overweight"

Use synonyms in a wide sense!

Phrases:

Words in templates can also be phrases. Example 1:

Required: [some certain ; people]

will match any question containing the phrases "some people" or "certain people".

Example 2:

Required: cannot [can ; t]

will match any question containing the phrases "cannot" or "can't".

Substitutes:

Certain sets of synonyms will be used in many different questions. Instead of repeating them in each template, they can be defined in the substitutes list, a list which is common to all templates in one language.

Example:

The substitutes list contains:

$buy=buy* purchas* [spend* # money] shop* sell* obtain* acquire* pay* squander* waste*

And this can then be used in a template such as

Required: $buy ; impulse* stimulanc* incite* desire* craving* longing* covet*

Intermediate words in phrases:

Phrases can also allow intermediate words. Example:

Required: [mental* psych* personalit* # ill* disord* sick*]

This will match not only "psychic disorder" but also "psychiatrically diagnosed disorder". Any number of intermediate words are accepted.

Specified intermediate words:

It is also possible to specify that certain intermediate words can or cannot occur. Example:

Required: [caus* : of # obesit*]

This will match both "cause of obesity" and "causing obesity" but not "cause of illness from obesity".

Multiple templates for one FAQ:

Sometimes, several different questions may match the same FAQ. You can therefore specify more than one template for each FAQ. Example:

Required: $eatingdisorder $obesity $slim $weight $eat ; cultur* newspaper* magazine* television* telly* cinema* movie* ad* placard* poster* bill civilization* civilisation* education* breed* manner* fashion* ideal* concept* / [am are is # not ; well]

FAQ Attributes:

Attribute	Description
ID:	Unique identifier to separate this from other FAQs.
Title:	A user-oriented textual description of the FAQ.
Question	One or more examples of questions in natural language.
Required:	Keywords which must exist in the user query for a match to be done.
Priority:	Keywords which will cause this FAQ to be shown before FAQs without this keyword.
Optional:	Keywords which may or may not exist in the user query. Words in "Priority" need not be repeated here.
Forbidden:	Keywords which may not occur in the user query. Use sparingly and carefully!
Non-envisaged words:	If the user query contains no more than this number of non-specified keywords, the response is shown to the user as a *probable match*. Otherwise, it is shown to the user as a *possible match*.
Body:	The HTML or XHTML text of the answer to the FAQ.

ID: ed-dia-bulimic-obese

Group: e-dia - Eating Disorders Diagnosis

Title: Are bulimic people obese?

Required: bulimic* [binge*; eat*] ; obes* overweight* fat* corpulent* stout* fleshy [weigh* # $much]

Optional: $people unusual*

Forbidden: anore*

Non-envisagaged words: 1

Priority: bulimi*

Body: Usually, the term <i>bulimic</i> is used for people with bulimic behaviour but who compensate for the overeating so as not to become overweight.

1. Try to think of every possible question which matches the answer.
2. Specify each matching question as a template.
3. Try to find synonyms in a wide sense. I have been using the Internet-based synonym dictionary for English at http://www.wordsmyth.net/. This is good because it also includes similar words.
4. Enter the classification, and try it out with different questions in the natural-language question-answering system to see that it works.

Required: decline* refus* abstain* rebuff* [pass* ; up] demur* [show* ; up] arriv* appear* present [drop* ; in] come came

Optional: thank* invit* accept* agree* consent* decid* $eatingdisorder promis* $people meet* welcom* greet* face* appoint* contact* [get `# together]

Limit: 2

"Why do people with eating disorders often accept an invitation, but then not show up?"

"Why does she refuse to show up when we agreed to meet?"

"Why did he not come when we had decided to get together at the Opera house?"

Required: $eat ; sensibl* rational* levelhead* reasonab* unreasonb* prudent* intelli* sane* insane* unrealist* realist* thoughtful* credib* understand* know* clearhead* bright* perspica* precept* astut* smart* apt suitab* witt* shrewd* $good together party* partie* band* company* bunch* group* gathering* alone lone* solitar* gregarious* secluded* single* desolate* separat* friend* accompan* unaccompan* [on ; $people ; own] [with by at in # $people]

Priority: $eatingdisorder

"Is it easier to eat sensibly together with other people?"

"Why do I binge eat more when I am alone?"

"Why am I more sensible at parties?"

Which food should I eat and which food should I avoid?:

Required: diet* weightplan [$eat $food nutrit* calory* calorie* ; plan control regulate control* regulat* keep* regulat* manag* restrain* guid* supervis* subdue* curb* hold* check* restrict* restrain* govern* harness* contain* confine* restrict* tame tune* modulat* adjust* keep* stead* stationar* balanc* stable* level stabilize* uniform* flat even constant*] / $eat ; $goodbad sensible right right* good* proper* [not ; too] correct* normal* sound* responsi* appropria* well suitabl* / $goodbad ; $food $slim /$obesity ; $cause

Priority: eat avoid choose select buy [$cause ; $obesity] right* well* correct* good*

Forbidden: [eating ; disorder] personalit* charact* temperament* psych* persona

Examples of questions which match this template:

- Help me develop an eating plan.
- How should I eat sensibly?
- What food is good and healthy?
- What can cause obesity?

The Forbidden clause above is included in order to avoid matches for questions like:

- Which personality factors can cause obesity?
- What is wrong with my psyche causing me to eat badly?

There are other, more appropriate answers to these questions.