**Argumentation Frameworks for Behaviour Change**

**Whats the difference from “normal” argumentation?**  
Our work focuses on argumentation in behaviour change. The difference is that in this setting the user *knows* that he or she engages in unhealthy behaviour and is *not* arguing in favour of that behaviour but provides arguments for his or her personal engagement in the behaviour in question.

For example, no rational person will argue in favour of smoking, eating unhealthy, not exercising, not saving water and electricity etc. However, many people still do one or more of the named things. They all have their personal reasons for doing so and our aim is to find out what those reasons are and the arguments that support them (and counter-arguments that attack them).

Example (taken from *Health Behaviour Change, Mason & Butler)*:

Nurse talking to smoking pregnant woman

Nurse: […] but you want to give this baby the best possible chance dont you?

Patient: Well, of course I do but Ive been a smoker for so long and it really helps me cope with the stress of balancing looking after my first child with going to work and so on […]

We can see that the patient is not defending *smoking* itself but only her actions (*it helps me to cope with stress*).

**What argumentation schemes will be used?**

The medical literature on behaviour change agrees that the success of behaviour change mostly depends on the motivation of the patient. Mason and Butler, however, find the word *motivation* too ambiguous and argue that if patients are asked how motivated they are, the answer will be an amalgam of *importance* and *confidence* and be less useful in understanding the patient’s position.

For example asking a patient *“How motivated are you from a scale of 0 to 10?”* might produce a *5.* Asking, on the other hand, two questions: *How important is it for you*? And *How confident are you?* Might produce 8 and 3 respectively. The 5 would therefore be the average of the two.

Mason and Butler define importance and confidence as: Someone is more likely to change a given behaviour if he or she believes that, on balance, it will lead to good outcomes (importance), and also believes that, across a range of situations, he or she will be able to master the skills needed to achieve that change (confidence).

They suggest the following guidelines for deciding whether to focus on importance or confidence:  
- if the importance level is depressed, focus on this if at all possible  
- focus on the lower number, importance or confidence, particularly if there is a large discrepancy between them  
- if they are approximately equal, start with importance

For our purposes, this devision seems practical since there would be no point arguing with a person, who does not believe changing their behaviour is important, that there are many ways he/she could change it. At the same time there is no point arguing with people about the importance of behaviour change if they already know it but do not know how to change

We therefore decided to use two argumentation schemes:

**1. Arguments from consequence (importance lower)**

For those who are not convinced that the behaviour change in question is important, we believe arguments from consequence are the best fit. The importance of a change in eating or exercise behaviour is weighed up against the importance of change in other, sometimes more important, matters such as saving money by buying cheap food, or being around the children in the evenings, rather than going to the gym (Mason & Butler). Under these circumstances value conflict with the practitioner is not uncommon. A practitioner might feel that it is very important to look after one’s health and prevent the onset of disease by adjusting one’s lifestyle. Many patients, however, feel otherwise, for a range of reasons. These differences in values will influence belief about health and illness and the perceived importance of a change in behaviour.

The person who is not convinced about the importance of the behaviour change might not be aware or not believe in the negative/positive consequences and therefore needs to be persuaded that it is important.

**2. Value-based practical reasoning (confidence lower)**

People who are aware of the importance of behaviour change often do not change their behaviour, despite knowing how important it is due to a conflict of values, as mentioned above. Under those circumstances we would like the patient through argumentation to arrive at a solution that does not compromise his/her values but also promotes the value of the behaviour change: *health.* This is comparable to *motivational* *interviewing* which is a goal-oriented, client-centered counseling style for eliciting behavior change by helping clients to explore and resolve ambivalence. Our goal would be to persuade the patient that his conflicting values (e.g. socialising and health) could be made congruent (Bardi & Schwartz), meaning that actions taken to express each of them are compatible with attaining the other value (e.g. socialising + health = going climbing with friends).

**A Chatbot for harvesting arguments**

**Long term goal:** Chatbot that is able to persuade people to engage in behaviour change  
**Medium term goal:** Chatbot that is able to discuss why people do not engage in behaviour change

**Short term goal:** Chatbot that is able to acquire arguments why people do not engage in behaviour change and counter-arguments for doing so

The first step in the development of a chatbot of such a kind is to populate a the knowledge base of the chatbot. Depending on whether the user believes that engaging in a particular change in behaviour is important or not, he or she will follow a certain path in the dialogue. The people who do not believe that behaviour change is important will be asked to name positive and negative consequences for pursing their values while the other group will be asked to name ways on how their values could be made congruent with the value “health”.

We will use “doing more exercise” as an example.

Definitions:

**Value:** The definition of “values” differ across the medical and psychological literature. We decided to adopt Schwartz’s (1994) definition: values are desirable transsituational goals, varying in importance, that serve as guiding principles in the life of a person.

Example: A person might prefer seeing friends over doing sports. The associated value with her behaviour could be “friendship”. The goal of her behaviour would be “to socialise” or “to maintain a friendship”.

Example 2: A person might prefer not going jogging in the park in the evening. The associated value with her “non-behaviour” would be “security”. The goal of her non-behaviour would be “to be safe/secure”

**Goal:** The goal that a certain behaviour (or not engaging in a certain behaviour) should achieve in relation to a value.

The chatbot values health so the way it handles dialogue has the built in assumption that is values health. But currently the aim is not to promote anything but only to inquiere but with the standpoint of valuing health

The ~~goal~~ of the chatbot is considered to promote the value ***health*** by the means of exercising more

The goal of the user is considered to act according to their values

The goal of the VBA-dialogue is to arrive at a common goal which promotes the user’s values as well as the chatbot’s values

how to define therse and know when conflict occurs because its very context dependent (comfortable running shoes?)

**Conflicting values**

Example: for a person who does not enjoy sport ***comfort*** and ***health*** (through more exercising) are conflicting values.

**Congruent values**

Example: for a person who enjoys spending time with her family ***fun*** and ***family*** are congruent values

**Value Hierarchy**

A hierarchy of values ordered by importance of the value to the person

Example (from most important to least important)  
family  
knowledge

comfort

health

fun

Research Question I: Can users link values to their arguments

Research Question II: Can users arrange their values given in the dialogue into a hierarchy

Research Question III: How do the argument graphs of VBA dialogues look like

Research Question IV: is there a pattern whether health is compatible with the values/value hieararchy of the user or not (CHAIN explain chain better)

what is the relationship between the theory of VBA (exrensions and so on) and the conclusions reached in dialogues

Note:

Influence of values on behaviour is not fully understood. We are not psychologists and do not make any assumptions on the psychological usefullness of this principle but solely focus on the argumentative aspect: assigning a value to an argument in order to reply with an appropriate counter-argument.

The same applies to the fact that we do not believe that a chatbot can replace a health professional. However, being able to identify the reasons why a person engages in certain unhealthy behaviour and being able to produce counter arguments that are consistent with the values and therefore triggering further reasons, is the first step into the right direction.

Example:

User: I dont like going to the gym because its expensive (Wealth/Money)

value-inconsistent argument: but you could do sports with your friends (Socialising/Friendship)

value-consistent argument: but you could go jogging in the park

One can see that both arguments are counter-arguments to an argument why someone does not do sports. But depending on what value the initial argument promotes, the counter-argument has to promote (or attack) the same value, otherwise it does not make sense.

**Example-Dialogues:**

Note: All interviewed participants don’tdo enough exercise

**User 1: low importance**

Chatbot: From a scale to 0 to 10 how important is changing your behaviour and doing more exercise for you?

User: 2

Chatbot: Okay, what is the main reason you do not do enough sports?

User: Because I do not have a lot of free time and in the little I have I want to socialise with my friendship

Chatbot: Understandable. What value is promoted by your behaviour – pick the one that describes it the best  
[presents list of values]

User: socialising  
Chatbot: Thanks! Now – can you think of any potential negative consequences of your current behaviour that promotes the value: socialising  
User: Well, I guess I drink too much and eat or do unhealthy things while socialising  
Chatbot: Okay, anything else?  
User: no

Chatbot: Thanks! What are the positive consequences of your current behaviour that promotes the value: Socialising

User: I have fun! I see my friends and we talk and catch up

Chatbot: I see, what else?  
User: I guess thats it…

Conducting interviews like shown above would allow us to populate a pro and con database for values, similar to the one below:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Socialising | | Family | | Comfort | | Knowledge | |
| Pro | Con | Pro | Con | Pro | Con | Pro | Con |
| - catch up and have fun with friends | - engaging in unhealthy activities | - spending time with my family | - spending the time infront of the TV and not moving | - relaxing after work | - not doing anything good for my body | - studying and getting a good grade in the exam | - neglecting my future health |

Note: for the value HEALTH – the pro and con could be filled in by a health professional

Those consequences can be used in later versions of the chatbot when conducting arguments-by-consequence-dialogues.

IDEA: Chatbot could ask in current version which argument the user ranks as “stronger”

or chatbot could ask – if youre aware of the bad conseqence, why are you engaging in the unhealthy behaviour, and therefore trigger more counter-arguments

User 2: low confidence

Chatbot: What is the main reason you dont exercise?

*User: “I dont like going to the gym, I feel uncomfortable and its often crowded at the times when I am free, like after work”*

Chatbot: what is the value your current behaviour promotes  
User: *dignity [*or *comfort]*

Chatbot: What advice would you give a friend that wants to do more sports and has the same value as you that would not conflict with her values

*User: “She could do sports outside of the gym, like running in the park or cycling home from work”*

Then the chatbot would ask why she is not doing exactly that. During the conversation it would become apparent whether the goals/values of the user are compatible with the chatbots goal – promote the value *health* through exercise

Compatible goals Incompatible goals

|  |  |
| --- | --- |
| [...]  Chatbot: So why dont you do that?  User: I am too scared to cycle in London (value: Security)  Chatbot: if your friend has the same problem – what would you advise her to do?  User: She could go running in the park when the weather is good and go swimming if its not, or just do some exercises at home  Chatbot: would you see yourself doing that?  User. Yes I guess I could | […]  Chatbot: So why dont you do that? User: The weather in london is too bad to cycle or jog (value: Fun/Joy)  Chatbot: if your friend has the same problem – what would you advise her to do?  User: Well she could go to the gym during lunch break or sign up for classes or join a sports team  Chatbot: would you see yourself doing that?  User: No gyms and sportsclubs are too overpriced in London, I would rather spend the money on something else (value: wealth)  Chatbot: Ok, so what other solution is there that would not compromise your values but allow you to exercise more?  User: I can’t think of any |

Conducting interviews like shown above would allow us to populate a database of value-tuples for HEALTH and the corresponding arguments, similar to the one below:

|  |  |  |  |
| --- | --- | --- | --- |
| Health vs socialising | | Health vs knowledge | |
| Arg | Counter-arg | Arg | Counter-arg |
| I want to spend time with my friends | You could go climbing together | I need to study for the exam | You could take regular 20-min study breaks and so a few exercises at home or go for a walk |
|  |  |  |  |
| Health vs security | | Health vs wealth | |
| I dont want to go jogging in the park in the evening | You could go swimming instead or do exercises at home | I would rather spend money on something else than the gym | You can go jogging for free |

**Research Question I:** Can users link values to their arguments  
**Hypothesis:** Every reason (argument) promotes a certain value which can be linked to the argument  
**Measurements:** compare whether users attached the same values to similar arguments

(note: should I include option – none?)

DO SURVEY TODAY – pairing values to args

**Research Question II:** Can users arrange their values given in the dialogue into a hierarchy  
**Hypothesis:** Users can arrange the values they have mentioned during the dialogue in a hierarchy. The first mentioned value is ranked the highest and the subsequently mentioned values are ranked lower  
**Measurement:** self-explanatory

for later use regarding the value chains/sequences whether the same hierarchies come to the conclusion thats health not compatible

**Research Question III:** How do the argument graphs of VBA dialogues look like

**Hypothesis:** Users will link other values to counter arguments of the counter arguments, different from the initial one

--→ WHAT TONY SAID

example:

argument: I want to spend time with friends (socialising)  
counter-argument: you could go climbing with them  
counter-argument: but the climbing wall is so far away and it would take time and money to get there (comfort)

**measurement:** construct argument graphs and compare them

**Research Question IV:** is there a pattern in the values/value hierarchy whether the chatbot’s goal is compatible with the user goal

**Hypothesis:** the order of the values named in the dialogue or the value hierarchy or just certain values can indicate whether the values of the goals of the user and the chatbot are compatible

example:

above:

comfort + security were congruent with health → can do more exercise and not conflict with values

comfort + fun + wealth were conflicting with health → cannot do more exercise without conflicting with values

**measurement:** create value-chains/hierarchies and the outcomes and compare them

SEQUENCE