Problem:

- people engage in unhealthy behaviour/activities. Examples: smoking, drinking, not exercising, not eating healthily etc.

- rational people are aware of the fact that what they are doing is not good for their health (in the long run). Not rational person will argue that smoking or not exercising is healthy. They know that. Henceforth, arguing with them about this is pointless.

- however, despite knowing that this behaviour is harmful for them, those people still engage in it because the engagement in it promotes values that they place higher than health

Solution:

*note: since we are not dealing with the actual persuasion-task yet, but only with the data acquisition, the agent does not give any counter-arguments or suggestions himself. The agent merely asks the user to think about counter-arguments and negative consequences himself.*

- find out whether the person’s goals are (hypothetically) compatible with the agent’s goal (health) with the help of practical reasoning with values

example: The person’s goal is is socialising (value: friendship). The agent’s goal is getting the person to exercise. Person could do sports with her friends (e.g. go climbing). Those goals are compatible

- if the goals are not compatible the agent should use *argumentation by consequence* to collect arguments why not exercising is bad (due to the negative consequences), as well as arguments against exercising due to the negative consequences this might cause upon achieving the agent’s goal.

Example: The person is a busy investment banker who barely sleeps and in the hours he doesnt sleep – he works. The agent figured out that none of the counter-arguments were successful and the goals are not compatible (his goals might be professionalism, wealth, dedication etc). The agent will therefore ask the person to list the bad consequences of not exercising as well as the bad consequences exercising might have on his values.

This would allow us to populate a database of arguments and counter arguments and their associated values. We might find patterns between compatible and incompatible goals and their associated values (e.g. *friendship* and *health* might always be seen as compatible by the person but *comfort* and *health* might not be). Additionally we would collect many bad consequences for not doing sports, as well as bad consequences associated with not pursuing other goals/values.

All those could later be used by a persuasion agent

research questions  
Q1 – is it possible with the help of an interrogative chatbot to find out whether the goals/values of a user are compatible or incompatible with the (chatbot’s) value *health*

Hypothesis 1 *–* assuming that the user is not lying and is confident about her values it should be possible to reach a conclusion in most of the cases

Q2 – what are there values that are mostly compatible with health and values that are not, as perceived by the user

Hypothesis 2 – there should be values that are mostly compatible with the value *health* and values that are mostly incompatible

For later experiment:

Q2.1: Would identifying a person’s values give us a clue about which argumentation schema a persuasion agent should use

Hypothesis 2.1 – following the argumentation schema depending on the user’s value hierarchy is conceived as more persuasive than the alternative

Q3 – how are the values that are given during the course of the dialogue arranged in a hierarchy?

Hypothesis 3: the value of the first argument is the value ranked the highest and the following ones are ranked lower

For later experiments – not sure yet, has to be rephrased:

Q4 – for incompatible goals: does thinking about arguments by negative consequence trigger a value shift? Is it possible to transform a health-incompatible goal into a health-compatible goal

Hypothesis 4: Making people think about negative consequences triggers a temporary value shift

methodology - experiments and evaluation of those

experiment design

A chatbot was set up that would ask the user for reasons why she sometimes/never exercised. The user would provide a reason (argument), for example

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

*T*hen the chatbot would ask what value the user associated with this argument that the user could pick from a list. In this case the user might choose *dignity* or *comfort*

Then the chatbot would ask what advice she would give a friend with the same values, how she could exercise more but not compromise the values. The person might reply

*“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 – more exercise (promote the value *health*)

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 |

|  |
| --- |
| Chatbot: Could you please list me some bad consequences of not exercising? Google if you need too  User: Coronary Heart Disease, strokes, high blood pressure, breathlessness, flabby body, little energy, stiff joints, osteoporosis, poor posture, overweight…  Chatbot: and what would be the bad consequences for you of you exercising under the certain conditions, taking your values into consideration  User. Well I would feel uncomfortable in a crowded gym, I could get a cold while cycling or jogging in the rain and I would feel unhappy |