Towards a Robotic Dietitian with Adaptive Linguistic Style

Hannes Ritschel, Kathrin Janowski, Andreas Seiderer and Elisabeth André
Human-Centered Multimedia, Augsburg University
{ritschel, janowski, seiderer, andre}@hcm-lab.de

Abstract

This work outlines a concept and the necessary building blocks for creating a persuasive and personalized robotic dietitian for everyday health-related support based on existing technology and recent research insights. Key components include natural language generation for the social robot's linguistic style, mobile sensing hardware for tracking nutrition, and machine learning for adaptation.

1 Introduction

In recent years, an increasing amount of health-oriented technology has come to the market, indicating a general trend of growing willingness and acceptance to use health-oriented mobile devices. This includes fitness trackers, smart watches and mobile tracking apps. While these devices keep track of the human's everyday nutrition, provide tips, reminders, or recognize anomalies in health-related behaviors, they typically use a traditional, touch-based Graphical User Interface (GUI) for interaction with the human. When it comes to diet support, mobile apps often record the intake of food with a GUI to calculate and recommend appropriate next dishes in a textual manner. In comparison, embodied agents, such as social robots, have the ability to provide more natural and multimodal interaction, including speech, gestures and facial expression. Therefore, research has investigated the use of robots as weight loss coaches, in the context of robot-assisted training and exercises, multimedia learning scenarios and education, as well as for long-term support of people with diabetes.

Recent research experiments offer a high potential to provide intelligent diet support, such as an adaptive robotic nutrition advisor, which aims to convince people to choose more healthy drinks with Reinforcement Learning (RL) [RSJ+18], the use of Natural Language Generation (NLG) for generating textual messages in a mobile dietitian app [AM18] or to give social robots the flexibility to adapt their linguistic style in terms of personality [RBA17], as well as mobile hardware to log the user's nutritional intake [SFA17]. When combined and embedded in the user's domestic environment, these technologies open up the possibility to sense the human's dietary intake and provide advice in a natural, multimodal and motivating manner.

In order to maximize a robotic dietitian's persuasiveness and to keep interaction interesting and engaging over a long period of time, respecting the user's individual preferences is important. For example, experiments investigating the similarity and complementary attraction effect report that the preferred and most effective robot's personality can depend on the user's own personality as well as on the task. Moreover, it has been shown that politeness impacts the perceived persuasiveness of recommendations by robotic elderly assistants [HLB+16]. In-situ studies indicate different politeness preferences for domestic social robots [RSJ+19], too. Since the robot's language is the primary communication modality, NLG and machine learning are key technologies to provide an individualized interaction experience and effective diet support. The following sections outline a concept and necessary building blocks to create a persuasive and personalized robotic dietitian based on the aforementioned technologies and observations.

In: Emilio Calvanese Strinati, Dimitris Charitos, Ioannis Chatzigiannakis, Paolo Ciampolini, Francesca Cuomo, Paolo Di Lorenzo, Damianos Gavalas, Sten Hanke, Andreas Komninos, Georgios Mylonas (eds.): Proceeding of the Poster and Workshop Sessions of AmI-2019, the 2019 European Conference on Ambient Intelligence. Published at http://ceur-ws.org

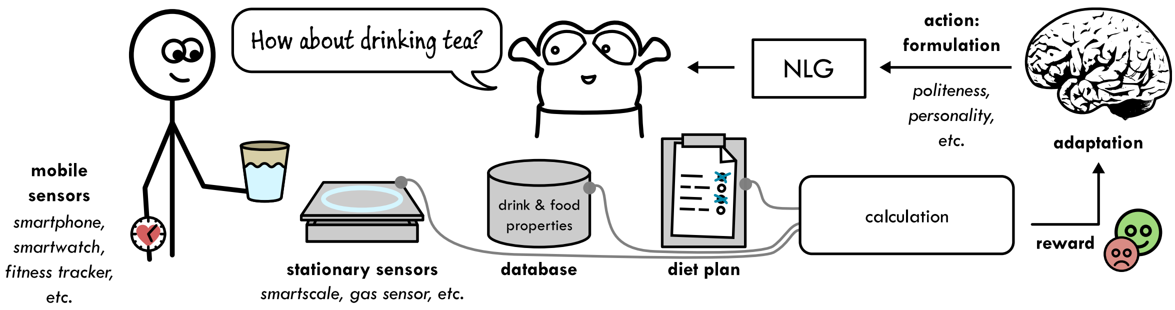

Figure 1: Overview of the proposed robotic dietitian

2 Adaptive Diet Support

Figure 1 illustrates the general idea. Information about the human's food consumption and activity throughout the day can be acquired with mobile and stationary hardware automatically, such as a smartphone, smartwatch or fitness tracker. When the user is on the go, he might use a traditional smartphone application's GUI or a virtual agent (e.g. an animated 3D model of the robot) for entering information and getting advice. Interaction in the user's domestic environment benefits from technology which can be installed stationary. This includes the social robot and additional sensors, such as a smartscale [SFA17, RSJ+18]. Both the problem of identifying the meal and estimating its amount needs to be solved before additional information from a nutrition database can be used to calculate the nutritional value and match the user's food consumption with the diet plan.

In order to provide personalized diet support we propose a machine learning approach: the learning agent's goal is that the user adheres to the diet plan, interpreting deviations as failure. Since we focus on the robot's linguistic style in this work, the robot can e.g. explore different politeness strategies to generate the most persuasive message for the individual user. RL is of special interest for this task since it allows to explore the robot's most efficient behavior autonomously. Based on the diet plan, acquired data from sensors, the user's activity and the meals' properties, a reward needs to be calculated. This positive or negative scalar indicates whether the robot's last action was expedient or not, so that the robot's linguistic style can be personalized to the user's reactions over time. Finally, the adaptation approach decides how to present the information. The NLG component generates the corresponding utterances and sends them to the robot. Additionally, multimodal cues, such as corresponding gaze behavior or facial expression can be added to emphasize the spoken language.

2.1 Nutrition Logging

Gathering information about the type and amount of consumed food is essential for the robot's advice and adaptation process. Other data, such as the calorie amount or the intake of specific nutrients can roughly be derived from this information with food databases. For behavioral analysis, the context in which food is consumed might be helpful (e.g. in the evening while watching TV), which might be extractable from smart home technology to a certain degree. The specific user's requirements based on demographic and health information (gender, age, weight, height, illnesses, medication) must be encoded by the diet plan. Additionally, the calorie consumption should be estimated e.g. by using data of a fitness tracker.

Several technical possibilities exist to sense the type of food. Most of them just work in specific use cases and usually a combination is required to allow a mostly complete automatic process. For example, image recognition is able to detect many types of food [MBO+18] as long as they are not pureed. In such cases gas sensors [DSA18] might be a better choice, nevertheless they can easily be disturbed by other odors. For both recognition methods the detection of self prepared food is a problem hard to solve if the preparation of the food was not logged. In such uncertain cases it might be a good choice to ask the user in the currently most convenient way as long as there is no perfect solution for unobtrusive, fast, mobile, automatic chemical food analysis available.

If the type of food is known, optical systems can roughly estimate the amount (weight), taking into account the vessel in which the food is located. One challenge is that usually not all parts of the food are visible to the camera depending on the perspective. A mobile scale [SFA17] can be a solution for this problem if higher precision is required, which however involves more effort than simply taking a photo.

One of the biggest problems is to detect the context in which a person eats or drinks. For behavioral analysis it might be sufficient to know how consciously a person eats or drinks the food as this is a major problem. Eye tracking is one technical option to give hints in this regard.

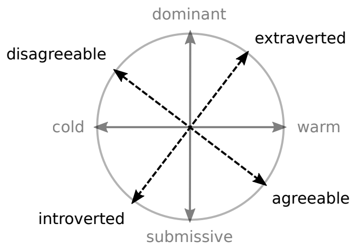

Figure 2: The axes defining the Interpersonal Circumplex. *Solid:* status and affiliation. *Dashed:* extraversion and agreeableness.

Phrasing	Affiliation	Status
"Drink tea."	cold	dominant
"How about drinking tea?"	cold	submissive
"We should drink tea."	warm	dominant
"You would probably like to drink tea."	warm	submissive

Figure 3: Recommendation phrasings using different politeness strategies. Warmth corresponds to positive politeness while submissiveness corresponds to negative politeness.

2.2 Personality and Politeness

After automatically sensing the user's nutrition, generating persuasive and personalized advice is the next task of the robotic dietitian. Apart from the actual message content, the way in which it is formulated and presented to the human plays an important role. Its expressed personality can be reflected in its multimodal output.

Interaction behavior is typically classified using the Interpersonal Circumplex [DWQP13]. It is defined by the two dimensions *status* and *affiliation*, with the former ranging from *submissive* to *dominant* and the latter from *cold* to *warm*. Alternatively, the same relationships can be expressed through the personality traits *extraversion* and *agreeableness*, which can be found at approximately 20 to 45 degrees relative to the other pair of axes [DWQP13]. Extraversion thus corresponds to a combination of high status and high affiliation.

Oakman et al. [OGC03] suggest that the Interpersonal Circumplex dimensions are also related to the politeness theory by Brown and Levinson. The so-called *negative face* is a person's desire to have autonomy with regards to their actions, while *positive face* is the desire to have others approve of one's own goals. *Positive politeness*, which minimizes threats to somebody's positive face, can be mapped to the affiliation dimension while the presence or absence of *negative politeness* roughly corresponds to status.

When looking at the robot's linguistic style, these relationships imply that extraverted persons are less concerned with threats to another person's negative face, but more inclined to apply positive politeness strategies such as treating the other person as a member of the same group. Conversely, introverted persons are more distant and submissive, and therefore avoid threats to the other party's autonomy while being less likely to use positive politeness. Figure 3 compares different phrasings for a simple example suggestion with regard to the expressed status and affiliation. With the flexibility of NLG the robot's dietary advice can be tweaked and formulated to increase its persuasiveness. Adapting the robot's politeness has recently been explored for a domestic robotic companion in the context of health-related recommendations based on template-based utterances [RSJ+19]. In contrast, NLG is a promising option due to the complexity of the diet context at hand.

2.3 Adaptation Process

The machine learning approach uses insights about the user's actual nutrition in comparison to the diet plan to improve the robot's behavior. RL can be used as a framework for optimizing details in the robot's linguistic style. For example, the robot's expressed politeness can be modeled as a nonstationary k-armed bandit problem [SB18], which is a reduced form of RL. The goal is to find the most effective of k actions $\mathcal{A}$ (politeness strategies) by estimating each action's value Q, which is calculated based on a scalar feedback, the so-called reward R. In each time step t the agent selects an action $A_t \in \mathcal{A}$, executes it, receives a reward R_t and updates the action's new value Q_{t+1} based on R_t, the old value Q_t and constant learning rate $\alpha \in [0, 1]$: $Q_{t+1}(A_t) = Q_t(A_t) + \alpha [R_t - Q_t(A_t)]$.

In order to react to changes in the user's preferences, Upper Confidence Bound (UCB) action selection [SB18] can be used for balancing *exploitation* and *exploration*, i.e., the agent's choice of the greedy (best) action with the highest Q-value versus exploring another supposedly suboptimal one. Therefore, $N_t(a)$ is the number of times action a already has been executed while $c > 0$ is a constant for controlling exploration. Based on this information the agent selects actions not only depending on their estimated values Q but also with regard to its uncertainty about the fact that their value might have changed in the meantime: $A_t = \arg\max_a \left[Q_t(a) + c\sqrt{\frac{\ln t}{N_t(a)}} \right]$.

By calculating the reward R based on the user's actual nutrition and the diet plan, the learning approach can optimize the robot's generated behavior over time by expressing itself in the most persuasive manner.

3 Conclusion

Our concept illustrates an approach for building a robotic dietitian, which personalizes its linguistic style to the individual user. With the ultimate goal of supporting the human's diet, persuasive messages are produced by a Natural Language Generation component, which enriches the robot's advice with personality-derived characteristics. Building on recent research, the reward for a reinforcement learning component is calculated based on the user's diet plan and its actual nutrition, making it possible to optimize the robot's messages for the individual user. We outlined necessary technologies to track the user's nutrition based on mobile and stationary sensor technology in an intelligent environment. All in all, we expect the robot to become more persuasive over time and thus foster a healthy lifestyle.

Acknowledgments

This research was funded by the Bavarian State Ministry for Education, Science and the Arts (StMWFK) as part of the ForGenderCare research association.

References

[AM18] Luca Anselma and Alessandro Mazzei. Designing and testing the messages produced by a virtual dietitian. In *Proceedings of the 11th International Conference on Natural Language Generation*, pages 244–253. Association for Computational Linguistics, 2018.

[DSA18] Chi Tai Dang, Andreas Seiderer, and Elisabeth André. Theodor: A step towards smart home applications with electronic noses. In *Proceedings of the 5th international Workshop on Sensor-based Activity Recognition and Interaction, iWOAR 2018*, pages 11:1–11:7. ACM, 2018.

[DWQP13] Colin G. DeYoung, Yanna J. Weisberg, Lena C. Quilty, and Jordan B. Peterson. Unifying the Aspects of the Big Five, the Interpersonal Circumplex, and Trait Affiliation. *Journal of Personality*, 81(5):465–475, 2013.

[HLB+16] Stephan Hammer, Birgit Lugrin, Sergey Bogomolov, Kathrin Janowski, and Elisabeth André. Investigating politeness strategies and their persuasiveness for a robotic elderly assistant. In *Persuasive Technology - 11th International Conference, PERSUASIVE 2016, Salzburg, Austria, April 5-7, 2016, Proceedings*, volume 9638 of *Lecture Notes in Computer Science*, pages 315–326. Springer, 2016.

[MBO+18] Javier Marín, Aritro Biswas, Ferda Ofli, Nicholas Hynes, Amaia Salvador, Yusuf Aytar, Ingmar Weber, and Antonio Torralba. Recipe1m: A dataset for learning cross-modal embeddings for cooking recipes and food images. *CoRR*, abs/1810.06553, 2018.

[OGC03] Jonathan Oakman, Shannon Gifford, and Natasha Chlebowsky. A multilevel analysis of the interpersonal behavior of socially anxious people. *Journal of Personality*, 71(3):397–434, 2003.

[RBA17] Hannes Ritschel, Tobias Baur, and Elisabeth André. Adapting a robot's linguistic style based on socially-aware reinforcement learning. In *26th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN 2017*, pages 378–384. IEEE, 2017.

[RSJ+18] Hannes Ritschel, Andreas Seiderer, Kathrin Janowski, Ilhan Aslan, and Elisabeth André. Drink-o-mender: An adaptive robotic drink adviser. In *Proceedings of the 3rd International Workshop on Multisensory Approaches to Human-Food Interaction*, pages 3:1–3:8. ACM, 2018.

[RSJ+19] Hannes Ritschel, Andreas Seiderer, Kathrin Janowski, Stefan Wagner, and Elisabeth André. Adaptive linguistic style for an assistive robotic health companion based on explicit human feedback. In *Proceedings of the 12th ACM International Conference on PErvasive Technologies Related to Assistive Environments*, pages 247–255. ACM, 2019.

[SB18] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning - An Introduction (Second Edition)*. Adaptive Computation and Machine Learning. MIT Press, 2018.

[SFA17] Andreas Seiderer, Simon Flutura, and Elisabeth André. Development of a mobile multi-device nutrition logger. In *Proceedings of the 2nd ACM SIGCHI International Workshop on Multisensory Approaches to Human-Food Interaction, MHFI@ICMI 2017*, pages 5–12. ACM, 2017.