A2HC 2021

XIII Workshop on Agents Applied in Healthcare

Part of the 20th International Conference on Autonomous Agents and Multiagent Systems

Organisers:
Dr Stefano Mariani, University of Modena and Reggio Emilia, Italy.
stefano.mariani@unimore.it
https://smarianimore.github.io/
Dr Sara Montagna, University of Bologna, Italy.
sara.montagna@unibo.it
http://apice.unibo.it/xwiki/bin/view/SaraMontagna/WebHome
Prof. Michael I. Schumacher, University of Applied Sciences, Switzerland.
michael.schumacher@hevs.ch
https://www.hevs.ch/en/minisites/projects-products/aislab/collaborateurs/uas-professor/schumacher-1800

Premise

Due to the ongoing global COVID-19 situation, **the workshop is asynchronous**, which means that each person interprets the schedule in their own time zone.
The aim of the schedule is:

- to get a commitment to spend time watching talks and engaging with comments
- to create synchronisation within each timezone, allowing for some discussion and social interaction

Organisation

As such, there is no static program of the workshop according to whatever time zone, as we are too dispersed to guarantee a nice experience to every attendee. Nevertheless:

- all the interested people, authors of the workshop or not, may exploit this shared document to **query presenters about their presentation**, either by using the "comment" functionality or by directly writing text in the appropriate section below each presentation. In any case, *please "sign" your comment* either with your name or with your contact information, so that the presenter or author can respond fruitfully
- all the interested people may ask to organisers (please, *include all organisers in Cc*) to arrange **live Q&A sessions** with the authors of the workshop papers, that organisers will then attend as moderators. The schedule of this live Q&A sessions will be agreed upon case by case amongst the paper authors and the interested people

Presenters (and other co-authors in general) MUST COMMIT to **regularly monitor this shared document** and reply to comments as appropriate, as a form of social engagement.

SOCIAL TABLE :)

Feel free to post your comments, questions, and appreciations down below in the suggested format. Please, remember to “tag” yourself with your name, more a more “friendly” discussion :)

Presentations

Session 1

Cohort and Trajectory Analysis in Multi-Agent Support Systems for Cancer Survivors.

Gaetano Manzo (speaker)*, Davide Calvaresi, Jean-Paul Calbimonte, Oscar Jimenez Del Toro and Michael Schumacher.*

Presentation video: https://youtu.be/jusJDX0GDFs
Presentation slides: https://www.slideshare.net/gaetanomanzo/manzoa2hcaamas
Comments section:

1. [Paul Marrow] Interesting because fully exploits the potential of multiagent systems to deliver both to patients and to clinicians. Questions: what is the user experience from both perspectives (assuming it has been tested with subjects in addition to data)? How is ethical use of patient data safeguarded through full understanding of consent? A related solution has been developed at UCL for intervention in colon cancer but only from the perspective of assisting clinical decision-making, not from the perspective of the patient.
    a. [Gaetano Manzo] Hi Paul, thank you for your comments and questions. As the first approach, we deliver the data analysis results only to the clinicians, which with a dedicated interface can monitor patients trajectory. Currently, we are working also with behavioral data aiming to deliver ad-hoc questions via the chatbot to the patient.
    b. [Gaetano Manzo] Data protection is the number one priority for us working with highly sensitive data. Therefore, we use Pryv a ready-to-use middleware for personal data and consent management. Here the link: https://www.pryv.com/ [Paul Marrow] I will definitely look at this. [Gaetano Manzo] Feel free to ping us if you may have any questions about it.
    c. Thank you for your solution. For sure I will have a deeper look at it. [Paul Marrow] I have not been involved in the development of it but I know some of the academics involved from the perspective of the analysis of clinicians user

experience of it. In: ACM Trans. Healthcare 2020 (from memory). [Gaetano Manzo] Thanks for sharing it.

2. [Stefano Mariani] Dear authors, interesting approach to chronic disease management. I personally like the “trajectory” metaphor, as during treatment patients actually follow a trajectory in a multi-dimensional space of health-related data. I have a few questions: is the ML pipeline you built to for trajectory prediction and analysis also encapsulated via Docker? Is it developed ad-hoc (“by hand”) for the specific task at hand? can it (or some portions of it) be automated (e.g. as for hyperparameter tuning in autoML)? Thanks!
    a. [Gaetano Manzo] Hi Stefano, thank you for your questions and observations. Yes the pipeline is dockerized into a Docker container and therefore easy to deploy. I am currently trying to put into a container each part of the pipeline and orchestrate it with Docker composer.
    b. [Gaetano Manzo] Yes it is an ad-hoc pipeline developed in Pytorch. Particularly for the data-preprocessing given the different data sources.
    c. [Gaetano Manzo] We are trying to automate each step of the pipeline. However, linking with the previous question, for some tasks the approach changes based on the data that we receive. We are also building an API where a system though a REST request can have the cluster and the trajectory of the patient (you insert the data of the patient and the API will output the cluster to which the patient belongs and the trajectory of a specific feature e.g., recurrency probability of the cancer).
    d. [Stefano Mariani] Thanks Gaetano. I was curious indeed as the issue of reproducibility, transferability, and generally speaking applying best software engineering practice to ML pipelines is a hot topic since recent years, and something I’m interested too as a consequence of a past EU project I worked for. Good to know your are working in that direction :) [Gaetano Manzo] Thank you :)

Optimizing Adaptive Notifications in Mobile Health Interventions Systems: Reinforcement Learning from a Data-driven Behavioral Simulator.

Shihan Wang (speaker)*, Chao Zhang, Ben Kröse and Herke van Hoof.*

Presentation video: https://youtu.be/qP-szlFKZ6k
Presentation slides: https://a2hc2021.github.io/assets/A2HC_SW.pdf
Comments section:

1. [Michael Schumacher] This work is very interesting. During the presentation, I was wondering if you plan a real experiment to measure in reality the efficiency of your context aware RL ? At the end of the talk, you mention actually that you plan a feasibility study. Could you tell us more about how you plan it? Which added data will you take into consideration? How do you plan to involve the users? Do you have any regulatory issues concerning privacy and ethics?

    a. [Shihan] Thanks, Michael. Yes, a feasibility study has been performed with real target users. 7 users used a smartphone exercise app in a week and the notifications were sent by our RL approach. During the study, we not only quantitatively collected the reaction of participants after receiving the notifications (e.g. reaction time and performance of physical activity), but also explored the experience of users based on questionnaires and interviews. This study was approved by the ethics review board of Utrecht University (it is a joint work with public health scientists). A paper presenting the detailed findings of this feasibility study is under revision for a journal publication. I would be happy to share the paper once it is published. :)
    b. [Michael Schumacher] Yes, I would be very interested to read the journal paper about the feasibility study.
    c. [Shihan] Great, Michael. I will send you the link afterward (I think I have your email).
2. [Ajith Vemuri] Interesting approach. Could you please tell a little more about “Urge” in your model? How did you estimate this? And when simulating how did you vary this parameter to accommodate varied behaviors (different people might have different urges to take actions)
    a. [Shihan] Hi Aijth, thanks for the questions. Good questions. I will try my best to answer them (I have also forwarded them to our co-author, who is a psychological researcher and knows the answer better than me. He may add answers later.) 1) I did not have the time to explain ‘urge’ in the presentation. Its definition and theoretical base is explained in the paper. In short, ‘Urge’ represents the user’s urge to perform the target activity at each time step. Based on psychological theories, it is one of key determinants of human decisions. For instance, when a user recently has a run, it oughts to temporarily lower his or her urge to perform a run in a short time. 2) In this paper, we did not take various behaviors of users into consideration. In fact, we have thought about it by clustering the users based on their behaviors first, then learning different delivery policies from data of each user group. A similar work (learn cluster based RL policy) has been done by researchers from VU (Title ‘Personalization of Health Interventions Using Cluster-Based Reinforcement Learning’). In the end, we did not follow this direction as we would like to focus on solving the interaction burden in this paper first.
    b. [Chao] Hi Aijth, I will try to add a bit to what Shihan already explained to you. So *urge to run* is one of the two key behavioral mechanisms included in the user model in order to simulate how a normal person decides to run (the other one is the activation level of running in memory). We made a very simple assumption that after each run, a person’s urge to run decreases to 0 and will take some time to recover, with the speed quantified by the parameter urge recovery rate. In this paper, we assumed the same parameter value for all users and the value was determined by the psychology literature and also a verification study, where we tested what combination of parameter values result in realistic running behaviors. So indeed we did not take individual differences in this mechanism into account and this will require future work.
    c. [Ajith Vemuri] Got it, thank you, Shihan and Chao.
3. [Stefano Mariani] Dear authors, I particularly enjoy your usage of bayesian networks for representing the user model. Did you performed structural learning for learning

also the structure of such network or was it predefined as shows in the slides? Also: it would be interesting to also use causal bayesian networks to track whether the user actually follows the suggestion, to possibly conduct what-if like analysis based on the causal bayesian network learnt (e.g."if I had sent one more notification, would it improve?").

   a. [Shihan] Hi Stefano, thanks for your comments. Yes! It would be interesting to take causal relation into consideration. Have you done some work in this direction or any suggested papers? I am very interested in the causal inference, but have not much experience. For now, we define the dependencies of features by ourselves (together with our co-author Chao based on psychological theories).
   b. [Stefano Mariani] Thanks for your reply :) The literature is vast and I am no expert, but a good starting starting point (at least for me) has been https://ieeexplore.ieee.org/document/9363924 from where you can easily "snowball" using its excellent bibliographic references (covering, e.g., most of Judea Pearl's work)
   c. [Shihan] Thanks.

4. [Pei-Yu Chen] Hi, as you use the historical running data and the weather data to derive the probability, as well as psychological insights, I'm wondering what would happen if the relationship between these factors change over time. For example a user used to like to go running on a sunny day but now they grow to like running on cloudy days for some reasons. Would you need to retrain the model? How would you tackle this kind of problem when applying this kind of technology in real life?
   a. [Shihan] Hi Pei-Yu, thanks for your question. This is an interesting question. I would answer it from two aspects. First, in this paper, we learn a delivery strategy from the historical data and psychological insights, then plan to use this learned policy as an initial policy. In this way, we aim to have a warm-start agent for our real users. In this sense, yes, the learned initial policy cannot be changed as the historical data is stable. And in fact, if we learn this from large-scale historical data, we shall obtain a generalized policy instead of a personalized one. Second, in real life, this warm-start RL agent shall be able to adapt to users' changing needs, because it will continue adjusting/optimizing the delivery policy based on the upcoming feedback from each user. In this sense, the model is able to learn a personalized for each individual user if he or she can use this agent for a while (unfortunately, I have not seen much actual applications in this case because it usually takes quite some time). I hope I answer your question. :)
   b. [Pei-Yu Chen] Thank you for your answer, Shihan. So just to clarify, the initial policy is learned from the historical/psychological data. This is a general policy. And then (in real life) it would be tuned in to the individual user based on their feedback to eventually have personalized policy, is it correct? I'm also working on a similar topic regarding behaviour support via notifications, but from a different approach. We are trying to use the user's values and norms (and other things) with machine reasoning techniques, so not so much RL based or data driven. I'm still at an early stage, just reading about pros and cons of different approaches. Your work is interesting, let's keep in touch :)
   c. [Shihan] Yes, Pei-Yu. You understand my points perfectly and I would be happy to keep in touch. You can find my email address in the presentation

slides. We are also open for collaborations. Your work sounds interesting. May I ask what do you mean by 'norms' (or how you define it)? And can you give me an example technique for machine reasoning? (I have heard about this concept and I am very curious). Thanks.

d. [Chao] Hi Pei-Yu, I am the co-author of Shihan. As I have a more psychology background, I am very interested in how ideas in psychology can contribute to machine learning or predictive modeling. I have done some work on combining a computational model of habit formation and machine learning to predict behavior in behavior change processes (see https://arxiv.org/abs/2101.01637). I am very interested in your approach of considering user's values and norms in machine learning. Would be curious to know more if you can explain your approach a bit more concretely.

Decision support for a vaccination campaign based on an agent-based model.

Emilio Sulis (speaker) *and Pietro Terna.*

Presentation video: https://vimeo.com/544860567 (password: a2hc)
Presentation slides: https://a2hc2021.github.io/assets/A2HC-2021-sulis_terna.pdf
Comments section:

1. [Sara Montagna] Hi Emilio! It would be nice to include in your model also the other type of vaccine. If I am not wrong study 15 refers only on vaccines based on mrna, what about adenoviruses based vaccines? How does the dynamic of the model change?
    a. [Emilio] Yes, we decided to propose an "average" response to an "average" type of vaccine. We will take care of your suggestion, thanks!

2. [Sara Montagna] It would be very very nice to compare your simulation results with real data which, if I am not wrong, you did not. it would be nice to cooperate in this work, for instance by collecting data of the two regions we are living in, and validate the model against real data.
    a. [Emilio] Good idea, in fact the model works well, but it is better to propose a comparison with real data to be more reliable… I will share the suggestions to Pietro Terna, as he is most engaged in the experiments at the moment! For sure I would certainly appreciate your cooperation for validation!
    b. [Sara Montagna] yes please! it would be very nice to devise a collaboration in this context
3. [Alina Cărunta] Hi! I noticed that you have used NetLogo. There are also tools like AnyLogic which could be used. Why did you choose NetLogo instead of others?
    [Emilio] Several reasons: it's FOSS. You can run the model on the web via a web browser. It's a sort of benchmark in ABM community, well-know around the world. has many advantages so...but yes, the drawbacks are many: limited computational capabilities, the code becomes complex for such a large project, increasing the number of agents becomes difficult/impossible... We are thinking to redo it in Python with a code written ad hoc for ABM by Prof. Terna, but it is still a hard work, not trivial.
    We know Anylogic, but there are also Repast or Gama that could work well…

[Alina Cărunta] I used Anylogic (academia version) for a project and I was limited to a few types of agents, while with NetLogo there could be more. [Emilio] Thank you, I will share your suggestion with Prof.Terna (the co-author more engaged in NetLogo programming). Anyway, I never used Anylogic for Genetic Algorithms.

Session 2

Augmenting BDI agency with a Cognitive Service: general architecture and validation in healthcare domain.

Sara Montagna, Stefano Mariani (speaker) *and Emiliano Gamberini.*

Presentation video: https://youtu.be/giVDFfjlu3o
Presentation slides: https://a2hc2021.github.io/assets/A2HC2021-BDI-cognitive.pdf
Comments section:

1. [Michael Schumacher] This work is very interesting. I have two questions. First question: As I understood that the work was tested inlab and not integrated with the real system yet. This may be quite challenging to achieve from a regulatory point of view, as you may be considered as a medical device. Do you have already plans for this, including clinical tests?
    a. [Sara Montagna] Hi Michael! Good point! As soon as the software the impact the clinical routines it will enter in the iter of certifications. Quite challenging! As a first step I would be very happy if the results of the precoded rules, which implement medical protocols, can be improved if integrated with a sort of analysis obtained from collected (big?) data. We are proceeding a bit slowly recently because the trauma team is the same involved in ICU and is actually full time in the management of the pandemic.
2. [Michael Schumacher] Second question: you mention in the conclusion that you want to address explainable techniques. This seems indeed crucial. Could you comment more how you want to address this?
    a. [Stefano Mariani] Hi Michael, thanks for your interest :) While I let Sara reply on the first point (as she is much more knowledgeable with those process than I am), I'll try to answer to your on point concern. First of all we still need to comprehensively assess efficacy of natively-explainable models such as those currently used in the paper (decision trees, Linear SVC) based on a larger sample size. On these models we already started doing investigations, e.g., regarding feature importance, which is widely used for explainability, that we didn't put in the workshop version of the paper for the lack of available space. Besides this, I'm personally interested in building and using using argumentation graphs to let the cognitive agent and the cognitive service debate over suggestions to deliver to the clinician (or actions to take): in this case, the argumentation graph would be made up of the predictions and feature importance values for the cognitive service, and of the plan context and actions for the cognitive agent. Whoever has the strongest "support

arguments" wins the debate and delivers the suggestion. Human feedback can also be included in the loop to do the arbitration. One possibility which I know exists but I'm not still sure whether to try to apply or not is post-hoc explanations extracted from black-box models such as NN or deep learning models...I am not an expert of the literature but "in the shoes" of a clinician I would not understand them as easily as, let's say, a decision tree or an argumentation graph...but I'm open to suggestions :D

b. [Michael Schumacher] This, argumentation graphs seem very interesting. As you are working an emergency setting, I guess also that the time constraints may be also important. Also having more that one human in the loop, as often decisions are share between different healthcare professionals.

Multi Agent Architecture for Automated Health Coaching.

Ajith Vemuri (speaker)*, Keith Decker, Mathew Saponaro and Gregory Dominick.*

Presentation video: https://youtu.be/yrq7L2sq2C8
Presentation slides: https://a2hc2021.github.io/assets/A2HC-MA3HCFinal.pdf
Comments section:

1. [Michael Schumacher] Your work is very interesting. However, in my opinion the two experiments miss the main point of the goal of the coaches, as you should measure the responses to interventions and adherence of the users, and not only the software architecture issues, such as messages. I think that this point is crucial in the choices of the different architecture modules and algorithms.
    a. [Keith Decker] Yes, absolutely. Unfortunately, Covid. So we now have two systems we have been testing for a year waiting to actually field :-) [BeSmart has really been testing for 20 months!] We are resubmitting the IRB protocols for both systems and hope to run trails over the summer. We focussed this paper on the architecture because we re-used a lot of parts, and are actually looking at a third "remix" for another application with another sensor.
    b. [Michael Schumacher] Thanks for those precisions! I fully understand the situation :-)
2. [Michael Schumacher] I am also wondering how you can ensure transferability and also coordination. Actually, those two elements may have incredible influences on the performances of health coaches.
    a. [Keith] Matt Saponaro's thesis has a lot more on transferability; we are slowly getting those parts published [e.g. *Matthew Saponaro, Ajith Vemuri, Greg Dominick, and Keith Decker. 2021. Contextualization and individualization for just-in-time adaptive interventions to reduce sedentary behavior. Proceedings of the Conference on Health, Inference, and Learning. Association for Computing Machinery, New York, NY, USA, 246–256. DOI:https://doi.org/10.1145/3450439.3451874*]. The coordination part is less clear; the underlying bones of this approach follow ideas from RETSINA and DECAF, both of which had a lot more focus on coordination. Ajith will be demonstrating this more with his thesis work :-) [Michael Schumacher] Thanks for the reference that I will read. This looks very interesting.

3. [Shihan Wang] Thanks for this interesting work. I have a quick question regarding the difference between the two case studies (maybe I miss it). Could you please help highlight it? Thanks.
    a. [Ajith Vemuri] Thanks for the question, the first system BeSmart has specific goals the user should achieve (for instance 3 bouts of physical activity for 20 minutes within a week (Target Heart Range 100 - 120). BeSmart system helps users achieve these goals by interacting with users (sending motivational messages and goal tracking messages). The second system Walking in JITAI has no “specific goals”, the idea here is to increase MVPA and reduce sedentary time by catching users in the moment of their sedentary or physical activities (for instance, a users walks his dog in the evening, the system detects that the user is walking and prompts a message saying “It looks like you are walking, why don’t you walk longer”. To do this the system takes the current context of the user (location, weather, current_activty, calendar data, response to previous intervention and threshold)
    b. [Shihan Wang] Thanks, Aijth. I really like the idea of JITAI in your second case studies (you may notice that we also follow this concept to motivate our own study). I guess there will be more challenges in the multi-agent setting for JITAI health coaching. I am slowly exploring multi-agent RL at this moment but could not see much real applications (barely nothing). Did you meet some issues in your practical study that can be shared with me? Or any suggestions for multi-agent coaching systems?
    c. [Ajith Vemuri] Thanks Shihan, I did notice your work uses JITAI, I found it quite interesting. Regarding issues with deploying multi-agents with health coaching, we did not really have any major issues. In our application each agent is quite independent of the other so it was straight forward. But we had issues regarding activity detection when using Apple watches, as our application needs to detect the current activity of users in real time, we had some issues to accurately detect activities (especially distinguishing between sitting and stationary standing)
    d. [Shihan Wang] Collecting the real-time data and performing detection automatically is indeed a challenge (we developed our own app to collect the data). I know there are some machine learning approaches for the activity detection (e.g. a survey paper https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7742959&casa_token=4HVwSQu_FEIAAAAA:ULdDh02Yzny62gbItTwEMd3OAMZ-CkaINQi42YL_OaKSF1HiUfifjSIFXD08ziXcLuNDgzAT0Rc&tag=1). But I am not sure whether any techniques can be easily reused in your system and for your specific purpose.
    e. [Ajith Vemuri] Thanks a lot for the resources Shihan, I will check them out.

[

[Keith] Yes, the “Just-in-Time” idea is apparently very popular in the behavioral health literature, but has seldom been actually tested in the wild (possibly because of the need for real-time sensor access: something we can get from AppleWatch but not from Fitbit.).[Ahh, just watched Shihan’s video. very cool. will be great to read the paper.]

4. [Paul Marrow] Interesting and very timely when Covid19 has reduced the capability of many people to take up sufficient physical activity for a long time. (Just starting again in the UK from where I am writing). However, shouldn't "coaching" be interpreted in a broader manner when referring to training for physical activity? Maybe it is more difficult to deliver in such a broader manner through the devices you are evaluating your services on. But I can imagine my personal trainer (if I had an agent-based one) interacting in multiple ways, and also giving me multiple ways to ignore or mute it. But, keep up the work, because if it is delivered only to a few subjects it can be good for their health. I'm now going for a walk.
    a. [Ajith Vemuri] Yes, you are absolutely right about "coaching" to be interpreted in a broad manner. Health coaching is a very large problem which not only includes what "times" the system interventions should happen but with what kind of messages (what text content motivates a specific individual). We haven't yet solved the entire coaching problem but are progressing towards it but solving specific sub problems. Enjoy your walk :-)

5. [Chao Zhang] As someone working in the field of e-coaching and JITAIs but not being familiar with multi-agent systems, I am very interested in your approach. Could you elaborate at a high-level what are the strengths (and challenges) of taking the multi-agent approach to JITAIs, when compared with other personalization methods? Would be really curious to hear!
   [Ajith Vemuri] Hello Chao, one the biggest reasons to adopt multi-agents is its distributed nature of computation. "Agents" here need not just run on servers but can run on edge devices (like phones and wearables). In the context of e-coaching this gives two big advantages 1. Privacy of data, with multi-agent approach data can reside on user's private resources. 2. Scalability: As computation is distributed, multi-agent systems scale quite well. The multi-agent approach becomes quite challenging when agents are interconnected and talk to each other

Feedback & follow-up

Feel free to put below any observation, criticism, comment on this edition of the workshop, and to use this whole document (which will stay online indefinitely) for keeping in touch.