Coaching Agent: Making Recommendations for Behavior Change. A Case Study on Improving Eating Habits

Supplementary Material

Jules Vandeputte
UMR MIA-Paris, AgroParisTech, INRAe, Université Paris-Saclay
Paris, France
jules.vandeputte@agroparistech.fr

Antoine Cornuéjols
UMR MIA-Paris, AgroParisTech, INRAe, Université Paris-Saclay
Paris, France
antoine.cornuejols@agroparistech.fr

Nicolas Darcel
UMR PNCA, AgroParisTech, INRAe, Université Paris-Saclay
Paris, France
nicolas.darcel@agroparistech.fr

Fabien Delaere
Danone Nutricia Research
Palaiseau, France
Fabien.DELAERE@danone.com

Christine Martin
UMR MIA-Paris, AgroParisTech, INRAe, Université Paris-Saclay
Paris, France
christine.martin@agroparistech.fr

ABSTRACT

This document provides the supplementary material for the article *Coaching Agent: Making Recommendations for Behavior Change. A Case Study on Improving Eating Habits* accepted for publication in Proceedings of the 21th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022)

ACM Reference Format:
Jules Vandeputte, Antoine Cornuéjols, Nicolas Darcel, Fabien Delaere, and Christine Martin. 2022. Coaching Agent: Making Recommendations for Behavior Change. A Case Study on Improving Eating Habits: Supplementary Material. In *Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), Online, May 9–13, 2022*, IFAAMAS, 4 pages.

1 RESULTS OBTAINED WITH DIFFERENT USER LEARNING RATES

As in the main paper, the results presented here were obtained by performing simulations of interactions between a coach and simulated users. The latter are still characterized by the same acceptability matrix $\mathbf{M}$ and starting preference vector Π_0 over the available items. The only difference here is the propensity to learn λ : we report the results of simulations for two values, $\lambda = 0.5$ and $\lambda = 0.9$. We consider the same set of user profiles, and the same set of strategies and initialization settings we reported in the main paper. Finally, we present the results for $N = 2000$ interactions, and simulations are made over 200 users.

Dependence over λ

Table 1, figure 1 and figure 2 present results for $\lambda = 0.5$, while table 2, figure 3 and figure 4 present results for $\lambda = 0.9$. Table 1 and Table 2 report, for each situation, the mean value of $\mathcal{V}(\Pi_T) - \mathcal{V}(\Pi_0)$ and the standard deviation for 200 simulations and $T = 2000$, respectively for $\lambda = 0.5$ and $\lambda = 0.9$. The results obtained, compared with results of the main paper for $\lambda = 0.2$, lead to some remarks. First, it is noticeable that for every strategy other than uninformed Q-learning, a higher user learning rate leads to better performance. This is not surprising, considering that a higher λ leads to a faster learning from the user, allowing him/her to perform more important behaviour changes in the same time T. While most of strategies are able to quickly adapt and make useful recommendations, Q-learning which needs much more exploration, leads the user to even worst behaviour when facing good tier users. Second, except for GA that appear to be slightly better than GES on good tier user with $\lambda = 0.9$, the relative performance of strategies is mainly maintained between set-ups and is the same as for $\lambda = 0.2$. We can therefore conclude that the performance of strategies is robust to λ. Third, regarding the standard deviation, one can note that the higher the λ, the greater the uncertainty. It is also noticeable that this effect is observable only for uninformed strategies (GES, GA, IBCFs, IBCF, GS and QL), but not for pre-informed strategies (iGES and tQL-10000). This can be explained by the fact that for users with high values of λ an accepted recommendation will have much more impact on $\mathcal{V}(\Pi_t)$. In fact as they modify more rapidly their habits, an accepted recommendation will lead to greater changes in Π_t, and mechanically on $\mathcal{V}(\Pi_t)$ given that the latter is an expected gain calculated from Π_t. While this will lead to low consequences on pre-informed strategies that propose accurate recommendation, the effect is much larger on uninformed strategies, because in these cases the coach is still learning and making recommendation errors.

Figure 1 and figure 3 are reporting the evolution over time $\mathcal{V}(\Pi_t)$, for respectively $\lambda = 0.5$ and $\lambda = 0.9$. Regarding the dynamic of $\mathcal{V}(\Pi_t)$ for the plotted strategies one can note that the main effect of λ is on the convergence speed. In fact, the higher the value of λ, the faster the convergence

Regarding $\mathcal{V}(\Pi_t)$ on figure 2 and figure 4 one can note that the final performance of iGES for both $\lambda = 0.5$ and $\lambda = 0.9$ is nearly the same. However, an higher learning rate leads to a faster convergence. On the other hand, results for both tQ-learning-5000 and tQ-learning-10000 are better for $\lambda = 0.9$ than for $\lambda = 0.5$, which indicates that the maximum performance of tQ-learning is not reached yet. This is confirmed by the curves of recommendation rate, which stays high even for $\lambda = 0.9$. These results confirm the advantage

Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), P. Faliszewski, V. Mascardi, C. Pelachaud, M.E. Taylor (eds.), May 9–13, 2022, Online.

of non-myopic strategies in finding alternatives trajectories in the space of probability vectors Π_t, preventing the user to get stuck in a non-improvable behaviour.

Conclusion

Overall the results presented in this supplementary material, for different user learning rates, confirm conclusions of the paper, drawn from results obtained with $\lambda = 0.2$. The effects of the parameters discussed in the main paper as the type of strategy, the starting habits of the user, the importance of prior knowledge and the non-myopic nature of strategies are still observable for the tested λ values. Overall it appears that, as expected, the main effect of λ is on the evolution speed of $\mathcal{V}(\Pi_t)$.

Consumer prototype		GES	GA	IBCFs	IBCF	GS	QL	iGES	tQL-10000
Good tier	μ	3.50	3.40	1.97	2.08	0.99	-0.67	4.67	**5.22**
	σ	1.56	1.39	1.77	1.86	1.71	1.44	1.12	0.84
Bad tier	μ	4.38	4.80	2.24	2.50	0.57	0.56	8.02	**9.57**
	σ	3.37	2.59	3.29	3.06	2.97	2.54	2.33	2.58

Table 1: Table of the mean μ and standard deviation σ of the expected score : $\mathcal{V}(\Pi_{T=2000}) - \mathcal{V}(\Pi_0)$, for *Good tier* and *Bad tier* consumers with learning rate $\lambda = 0.5$ depending on the coaching strategy.

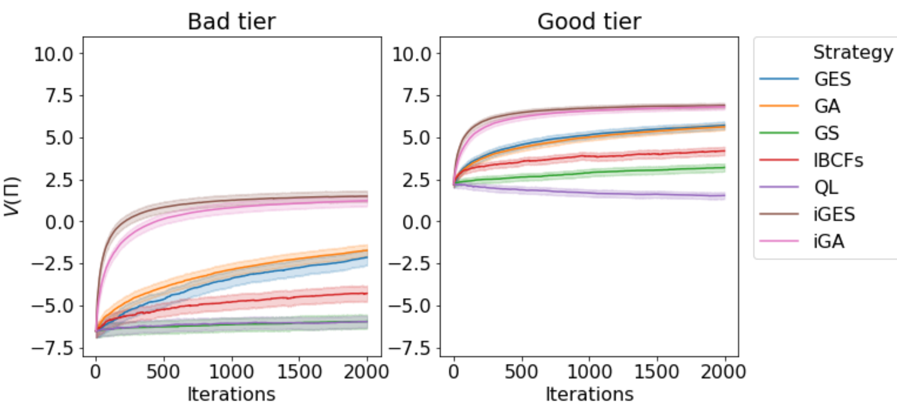

Figure 1: Comparison of $\mathcal{V}(\Pi_t)$, $(0 \leq t \leq T = 2000)$ for two informed strategies (iGA and iGES) and five uninformed strategies (GA, GES, IBCFs, GS and QL) for both *Bad tier* and *Good tier* prototype users with learning rate $\lambda = 0.5$. The colored area around the curves represent the 95% confidence interval.

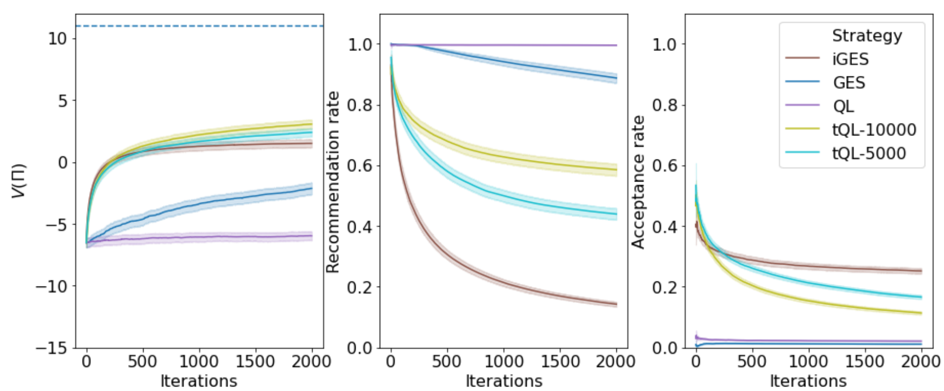

Figure 2: Comparison over 2000 interactions of $\mathcal{V}(\Pi_t)$, $(0 \leq t \leq T = 2000)$, the recommendation rate and the acceptance rate for GES, Q-Learning, iGES and trained Q-Learning on 5,000 and 10,000 steps, for *Bad tier* users with learning rate $\lambda = 0.5$. The colored area around the curves represent the 95% confidence interval.

Consumer prototype		GES	GA	IBCFs	IBCF	GS	QL	iGES	tQL-10000
Good tier	μ	4.20	4.39	2.05	2.40	1.41	-1.40	4.82	**5.83**
	σ	1.70	1.38	2.56	2.63	2.12	2.20	1.12	0.80
Bad tier	μ	5.94	6.67	2.50	2.83	0.86	0.82	8.19	**10.66**
	σ	4.19	3.20	3.62	3.97	3.29	2.86	2.33	2.79

Table 2: Table of the mean μ and standard deviation σ of the expected score : $\mathcal{V}(\Pi_{T=2000}) - \mathcal{V}(\Pi_0)$, for *Good tier* and *Bad tier* consumers with learning rate $\lambda = 0.9$ depending on the coaching strategy.

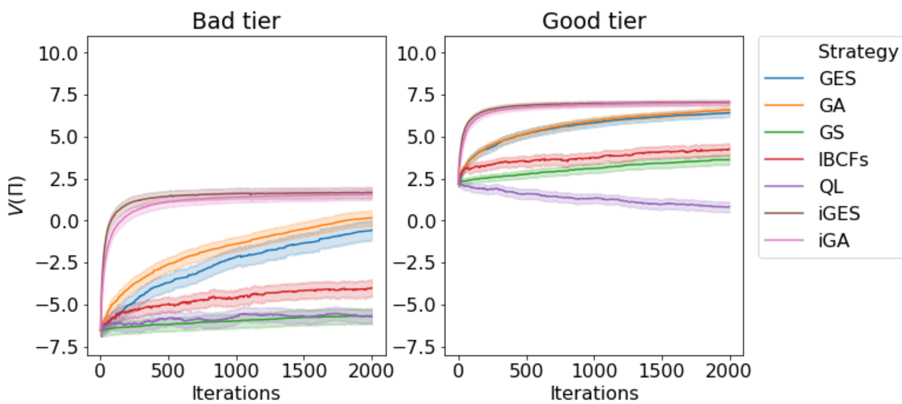

Figure 3: Comparison of $\mathcal{V}(\Pi_t),\ (0 \leq t \leq T = 2000)$ for two informed strategies (iGA and iGES) and five uninformed strategies (GA, GES, IBCFs, GS and QL) for both *Bad tier* and *Good tier* prototype users with learning rate $\lambda = 0.9$. The colored area around the curves represent the 95% confidence interval.

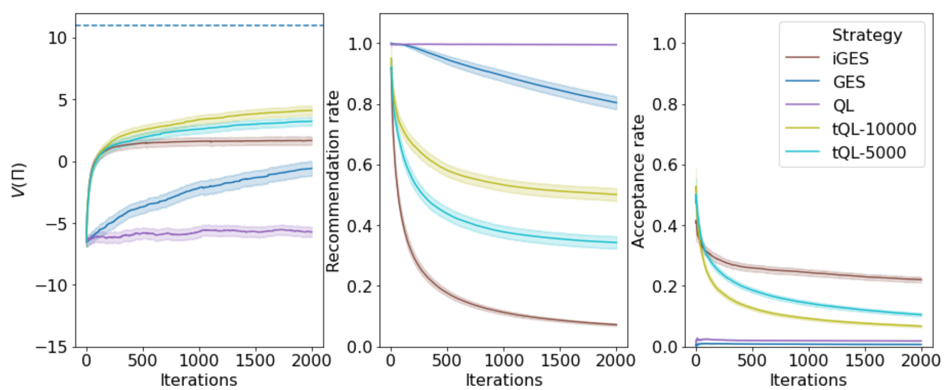

Figure 4: Comparison over 2000 interactions of $\mathcal{V}(\Pi_t),\ (0 \leq t \leq T = 2000)$, the recommendation rate and the acceptance rate for GES, Q-Learning, iGES and trained Q-Learning on 5,000 and 10,000 steps, for *Bad tier* users with learning rate $\lambda = 0.9$. The colored area around the curves represent the 95% confidence interval.