

生活習慣改善を支援する強化学習に基づく最適介入手法の 初期評価～目標時刻に就寝させることができるのか～

高橋 公海^{1,a)} 幸島 匡宏^{1,b)} 橋本 遼^{1,c)} 三宅 明日香^{1,d)} 倉島 健^{1,e)} 戸田 浩之^{1,f)}

概要：生活習慣病の増加は社会的な課題となっており、規則正しい食生活や十分な睡眠を確保することが予防に効果的であると知られている。我々は、ユーザが目指す生活リズム（ある所望の時刻に食事をする、就寝するなど）の実現を目的として、強化学習を用いてユーザが無理なく目標達成出来るよう支援する技術の開発に取り組んできた。本研究では、睡眠に関する生活リズムの改善を題材として、我々が構築した技術を用いてユーザが目標とする時刻に就寝することを支援できるか検証する。具体的には、生成したおすすめスケジュールを slack を用いて通知するシステムを実装し、評価実験を行った結果について報告する。

1. はじめに

生活習慣病の増加は社会的な課題となっており、予防のためには規則正しい食生活や適度な運動、十分な睡眠を確保することが重要であると言われている [1]。健康的な習慣形成を支援するために情報提示や介入を行うウェアラブルデバイスも身近なものとなっており、例えば Fitbit ^{*1}では1時間の歩数が少ない場合に運動を促すリマインダや設定された時刻に睡眠を促すリマインダ、Apple watch ^{*2}にはリラックスするよう深呼吸を促す呼吸リマインダなどがある。しかしながら、ユーザに行動を促す（介入する）これらのリマインダは比較的シンプルなルールに基づいており、上手くいかない例もある。

例えば、ユーザが十分な睡眠時間を確保するために、いつもより1時間早く就寝したい場合について考える（図1(a)）。「夜11時に寝る」ことをユーザが目標とした場合、目標とする時間の少し前にリマインダを設定し、寝る時間を知らせるといった方法が考えられる。しかし、普段の生活パターンのまま寝る時間だけを早くしようと試みても、通知を受けた時点でいつも寝る前に行っている一連の行動が終わっていない等の理由で、通知に従って行動することが難

しく上手くいかない（図1(b)）。効果的な介入を行うためには、ユーザが目標としている行動だけを変えようと介入するのではなく、目標達成に影響する他の行動についても考慮する必要がある。

我々はこれまで、ユーザが目指す生活リズム（ある所望の時刻に食事をする、就寝するなど）の実現を目的として、強化学習 [2] を用いてユーザが無理なく目標達成出来るよう支援する技術の開発に取り組んできた [3]。図1(c)は提案手法の適用例を示したものであり、ユーザに18時頃夕食を食べよう促すことでユーザの将来の行動を変え（いつもより早い時間に入浴、午後11時に寝るという目標を達成出来るように支援する。このように、ユーザの生活パターンと介入が受け入れられる可能性を考慮し、目標（ユーザが設定した時間と行動）を達成するためにいつどの行動を促すように介入すると良いか、という方策を学習する手法を構築した。我々の手法は図1(b)のようなりマインドよりも有効であることを既にシミュレーションで確認している。

本研究の目的は、我々が構築した技術をシステムに実装し、ユーザ実験を通じて効果の検証を行うこと。また、今後システムを洗練させるための知見を得ることである。そのために構築したシステムの概要を図2に示す。まず、ユーザは毎日の行動をライフログとしてカレンダーアプリケーション ^{*3}に入力する。次に、システムは事前にユーザが設定した目標（時間、行動）の情報とライフログデータを利用して、目標達成のために最適な方策を得る。方策とは、時間帯（0～23時）×ユーザが行なっている行動×介入

¹ NTT Service Evolution Laboratories, NTT Corporation
1-1 Hikarinooka, Yokosuka-shi, Kanagawa-ken, 239-0847
Japan

a) masami.takahashi.xh@hco.ntt.co.jp

b) masahiro.kohjima.ev@hco.ntt.co.jp

c) ryou.hashimoto.dr@hco.ntt.co.jp

d) asuka.miyake.zw@hco.ntt.co.jp

e) takeshi.kurashima.uf@hco.ntt.co.jp

f) hiroyuki.toda.xb@hco.ntt.co.jp

*1 <https://www.fitbit.com/>

*2 <https://www.apple.com/watch/>

*3 <https://calendar.google.com/>

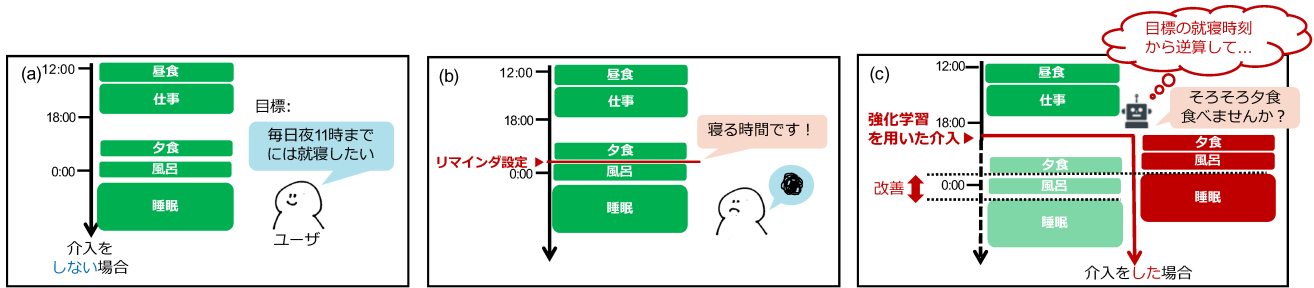


図 1: ユーザがいつもより早く就寝することを目標とした例. (a) ユーザが「夜 11 時、就寝」と目標を設定する. (b) 23 時にアラームを設定するようなシンプルな介入の場合、普段寝る前に行う行動を終えていなければ、介入に従って行動することが難しい. (c) は目標とする時間に就寝できるように逆算して、寝る前に行う一連の行動を早めるようユーザに促して目標達成を支援する.

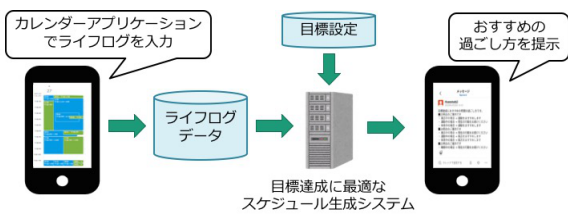


図 2: 実験用システム概要

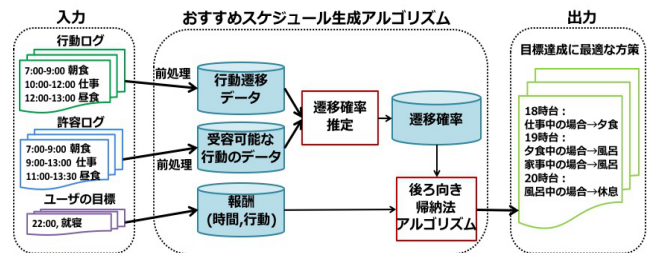


図 3: おすすめスケジュール生成システム

しておすすめする行動から構成される介入プランを指し、前日分までのライフログデータを使って 24 時間分の方策を算出する。最後に、方策をもとにユーザに提示するためのおすすめスケジュールの文面を作成し、slack^{*4}でユーザに通知するというものである。

今回の実験では睡眠に関する生活リズムの改善を題材として、ユーザが目標とする時刻に就寝することを支援できるか検証する。10 名の参加者を対象とし、あらかじめ実験期間中に各自が取り組む就寝時刻の目標設定を行ってもらった。実験期間は 4 週間で毎日ライフログ入力を依頼し、我々の手法で生成したおすすめスケジュールをシステムから通知される介入あり期間が 2 週間、何も通知されない介入なし期間を 2 週間設けた。評価は事前・事後のアンケート結果比較、および実験期間中に毎日取得していたアンケート結果を利用して定量的に分析し、おすすめスケジュール提示の有無が目標達成度合いにどう影響したか、また今後システムを改良していくための方向性を明らかにする。

2. 関連研究

健康行動を促すアプリケーションやデバイスは様々なものがあり、活動量・心拍・体温・呼吸数などを測定可能なセンサを搭載したウェアラブルデバイスは身近なものとなっている。これらはユーザの健康に関するデータを収集し、情報を通知することを容易にした。健康行動を促すための介入方法については様々な研究が行われており、Daskalova

ら [4] は、ウェアラブルデバイスで蓄積した睡眠データの分析やコホート研究に基づく睡眠習慣を改善するための情報推薦を行っている。実験から得られた知見によると、推薦に従うことが難しかったユーザに共通する理由として、スケジュールに合わなかったことが挙げられている。我々はこの点に着目して研究を進めているが、我々の過去の研究 [3], [5] ではモデルベース強化学習を用いた手法の構築とシミュレーションでの評価に留まっていた。本研究は我々の手法をシステムとして実装し、ユーザ実験により実際に介入を行って効果の検証を行うものである。

3. おすすめスケジュール生成システム

おすすめスケジュール生成システムとは、ユーザの日々のライフログと目標 (時間、行動) を入力として、目標達成に向けて最適なおすすめスケジュールを出力するものである。システムの概要を図 3 に示す。システム内部のアルゴリズムに関しては先行研究 [3] で詳細を説明しているため、ここでは主に入力となるライフログ、アルゴリズムの概要、出力されるスケジュールについて説明する。

3.1 ライフログデータ

入力となるライフログデータは行動ログおよび許容ログから構成され、参加者は表 1 の State に示す 15 種類の行動を対象に記録を行う。図 4 は入力例を示している。記録内容は「行動の種類」「開始時刻」「終了時刻」であり、カレンダーアプリケーションに予定として登録する。時間の粒度は 15~30 分程度を目安とした。記録の際には、仕事を

*4 <https://slack.com/>

表 1: 状態, 行動 (介入), 報酬

State	1: 睡眠, 2: 朝食, 3: 昼食, 4: 夕食, 5: 間食, 6: 出社, 7: 仕事, 8: 退社, 9: 家事, 10: 運動, 11: 休息, 12: 風呂, 13: 趣味, 14: 飲酒, 15: 買い物
Action	1: 睡眠, 2: 朝食, 3: 昼食, 4: 夕食, 5: 間食, 6: 出社, 7: 仕事, 8: 退社, 9: 家事, 10: 運動, 11: 休息, 12: 風呂, 13: 趣味, 14: 飲酒, 15: 買い物, 16: 何も介入しない (介入無し)
Reward	ユーザの目標 (例: 22 時に寝る)

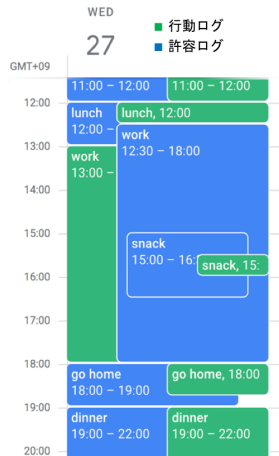


図 4: 行動ログと許容ログの入力例

しながら間食をとる, のように 1 つの時間帯に複数の行動が存在しても良いものとした. また, 対象としている行動以外を行っている時間帯は空白とすることも可能である.

行動ログは実際に行った行動を記録していくため比較的理解しやすいが, 許容ログについては日頃意識しない概念であり, 事前説明会及び実験の初期段階で考え方や入力時の注意事項など複数回フォローを行った. 許容ログを入力する際には次の 2 点がポイントとなる. まず, 仮に昼食を 12:00~12:30 の時間帯でとっていた時に, その日一日を振り返って「12:00~13:00 の時間幅の中であれば, 昼食をとる時間を動かすことが可能だったかもしれない」と考えられる場合は許容可能な時間を「12:00~13:00」と記録する (図 4). これにより, 我々は参加者が昼食をとる時間帯の変更を許容できる時間幅を知ることができる. また, その日行っていない行動 (例えば運動) であっても, 「19:00~19:30 の時間帯ならば運動することが可能だったかもしれない」と考えられる場合は「19:00~19:30」と記録する. 許容ログは 1 日の行動の記録を振り返り, 出来るだけ当日もしくは翌日に入力するよう依頼した.

3.2 強化学習による最適方策の獲得

ここでは図 3 のおすすめスケジュール生成アルゴリズムについて概要を説明する. 本問題では, 強化学習における状態・行動・報酬を, それぞれユーザの行動・アプリケーション

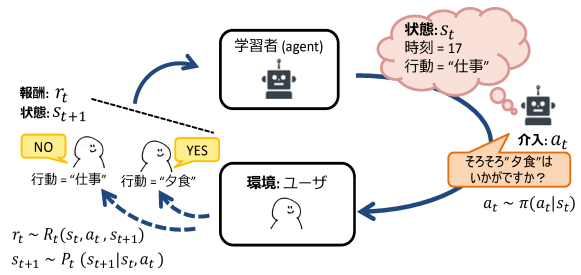


図 5: エージェントと環境 (ユーザ) とのインタラクション.

による介入・ユーザの目標に対応させる (図 5). 我々は以下の MDP (Markov decision process) の定義に基づいて, エージェントが最適方策を獲得できるよう手法を構築した. ユーザとエージェント間のインタラクションは有限期間非斉時的マルコフ決定過程 (finite-horizon time-inhomogeneous MDP) [6] でモデル化し, $\{S, A, \{\mathcal{P}_t\}_{t=0}^{T-1}, \{R_t\}_{t=0}^{T-1}\}$ の 4 つ組で定義する.

3.3 State

$S = \{1, 2, \dots, |S|\}$ は状態の有限集合である. 本研究における状態とはユーザが行っている行動を指し, 次の 15 種類を対象とした: 睡眠, 朝食, 昼食, 夕食, 間食, 出社, 仕事, 退社, 家事, 運動, 休息, 風呂, 趣味, 飲酒, 買い物 (表 1). 例えば, 「8 時 朝食」「10 時 仕事」「12 時 昼食」が状態となる.

3.3.1 Action

$A = \{1, 2, \dots, |A|\}$ は行動の集合であり, 本研究においてはユーザへの介入を行動とする. $|\cdot|$ は集合中の要素の数を示しており, T は行動の長さを示す. 行動空間 A は, エージェントがユーザに表 1 中にある Action のいずれかの行動を促す介入をする, もしくは介入を行わないという変数で構成される. i.e., $|A| = |S| + 1$. エージェントは介入を選択した後, ユーザに「夕食はいかがですか?」「寝る時間です」といったメッセージを提示することを想定している.

3.3.2 遷移確率

\mathcal{P}_t は遷移確率, $\mathcal{P}_t(j|i, k = j)$ は行動 k が t 番目の時間に行われた時に状態 i から状態 j へ遷移する確率を示している. これは, 介入有りもしくは無い場合のユーザの行動遷移を定義している. もし $\mathcal{P}_t(j|i, k = j)$ の値が大きければ, ユーザは j という行動を行うよう促すエージェントの介入に従う確率が高く, 次の時間に行動 j に遷移する. もしユーザへの介入が行われない場合には, ユーザの次の行動は $\mathcal{P}_t(j|i, k = 16: \text{none})$ によって決定する. このように非斉時的な MDP を用いることで, 朝や夜のように時間帯に依存したユーザの行動遷移をモデル化することができる. 遷移確率の推定はユーザの行動ログと許容ログを用いることで可能であり, 詳細は先行研究 [3] を参考にされたい.

3.3.3 Reward

\mathcal{R}_t は報酬関数であり、 $\mathcal{R}_t(i, k, j)$ は t 番目の時間に介入 k を行うことで状態 i から状態 j に遷移する時に得られる報酬である。既に述べたように、我々は時間に依存した遷移確率と報酬関数を用いる。ユーザが目標とする時間に目標の行動を実施出来た場合に、正の報酬を得るものとして設計する。本実験では、報酬関数 \mathcal{R} はユーザの理想とする目標（ユーザが設定した時間と行動）によって定義される。例えば、ユーザが十分な睡眠を確保するために目標とする状態 s_g を「睡眠」、 t_g を「22:00」と設定した場合について考える。報酬は基本的には、目標を達成した場合に大きな正の値 r_g ($r_g > 0$)、介入無しの場合に 0、それ以外の場合は負の値 r_{itv} ($r_{itv} < 0$) を与える*5 ($|r_g| \gg |r_{itv}|$)。つまり、次の状態 s_{t+1} が t_g (22:00) に目標とする状態 s_g (睡眠) であれば、報酬 r_g を得る (それ以外は r_{itv} を得る)。このように設計することで、エージェントは目標達成に繋がる時だけ介入を行い、それ以外は介入しないという振る舞いをする。そのため、不必要に何度もユーザに介入することを避けることが出来る。よって、報酬は下記のように設計する。 I は指示関数 (indicator function) である。

$$\mathcal{R}_t(s_t, a_t, s_{t+1}) = r_g I(s_{t+1} = s_g, t+1 = t_g) + r_{itv} I(a_t \neq \text{none}).$$

3.3.4 後ろ向き帰納法アルゴリズム

推定された遷移確率と報酬関数が与えられた時、我々の手法では最適な方策を求める。ここではエントロピー正則化強化学習 (RLER) [7] を用いるものとし、最適化の定義は RLER に従う。RLER の目的は、報酬和と方策エントロピーを最大化する最適な方策 π^* を求めることである。

$$\pi^* = \arg \max_{\pi} \mathbb{E}_{h_T}^{\pi} \left[\sum_{t=0}^{T-1} \{ \mathcal{R}(s_t, a_t, s_{t+1}) + \alpha \mathcal{H}(\pi_t(\cdot|s_t)) \} \right],$$

$\mathbb{E}_{h_T}^{\pi}$ は π に従うエピソード h_T の期待値、 $\mathcal{H}(\pi(\cdot|s_k))$ は分布 $\{\pi(k|s_t)\}_k$ のエントロピー、 α はエントロピー項の寄与度を制御するハイパーパラメータ*6である。エントロピー項は一様分布に近付くほど大きくなり、大きなエントロピー項を持つ場合に方策は決定論的ではなく確率的なものになる。報酬和の値が近い複数の介入行動が存在する時に確率的に行動が選択され、ユーザが介入に飽きてしまうことを避けることができる。そのため、最適な方策は大きな報酬を得ることができ、大抵の場合決定論的ではないことが期待できる。なお、この RLER はハイパーパラメータ $\alpha = 0$ の時に一般的な強化学習の枠組みと同一である。我々は RLER の価値関数を下記のように定義した。

*5 今回の実験では目標を達成した場合に大きな正の値 $r_g = 100$ 、何も介入を行わない場合は 0、それ以外の場合は $r_{itv} = -3$ と設計した。

*6 実験では $\alpha = 0.01$ を用いた。

Algorithm 1 後ろ向き帰納法アルゴリズム

Input: \mathcal{P} : transition probability, \mathcal{R} : reward function, α : hyperparameter
Output: $\{Q_t^*\}_t, \{V_t^*\}_t$: value function, $\{\pi_t^*\}_t$: policy
 1: Set $t \leftarrow T$ and $V_T(s) = 0$ for all $s \in \mathcal{S}$.
 2: Set $t \leftarrow t - 1$
 3: Compute $Q_t(s, a)$ following Eq. (1) for all $s \in \mathcal{S}$ and $a \in \mathcal{A}$.
 4: Compute $V_t(s)$ for all $s \in \mathcal{S}$ following Eq. (2).
 5: Compute $\pi_t(a|s)$ for all $s \in \mathcal{S}$ and $a \in \mathcal{A}$ following Eq. (3).
 6: If $t = 0$, stop. Otherwise, return to step 2.

$$Q_t^{\pi}(s, a) = \mathbb{E}_{h_T}^{\pi} \left[\sum_{t'=t}^{T-1} \{ \mathcal{R}_t(s_{t'}, a_{t'}, s_{t'+1}) + \alpha \mathcal{H}(\pi_{t'}(\cdot|s_{t'})) \} \middle| s_0 = s, a_0 = a \right].$$

また、RLER の最適なベルマン方程式は、

$$Q_t^{\pi^*}(s, a) = \mathbb{E}_{s' \sim \mathcal{P}_t(s'|s, a)} [\mathcal{R}_t(s, a, s') + V_{t+1}^{\pi^*}(s')], \quad (1)$$

$$\text{where } V_t^{\pi^*}(s) = \alpha \log \sum_{a'} \exp(\alpha^{-1} Q_t^{\pi^*}(s, a')). \quad (2)$$

この最適な価値関数と方策は Alg. 1 に示す後ろ向き帰納法アルゴリズムで求めることが出来る。RLER の最適な方策は次の確率的な方策で与えられる：

$$\pi_t^*(a|s) = \exp(\alpha^{-1} \{ Q_t^{\pi^*}(s, a) - V_t^{\pi^*}(s) \}). \quad (3)$$

3.4 おすすめスケジュール

システムから出力される方策は、時間帯 (0~23 時) × ユーザが行なっている行動 × 介入しておすすめする行動から構成される。イメージとしては、時間帯ごとに図 6 のような表を得ることが出来、値の大きい介入を選択することがエージェントの最適な戦略となる。今回のシステムではリアルタイムにユーザの行動を認識することが出来ないため、各時間帯毎にユーザが行なっている行動と、介入しておすすめする行動の組を「おすすめスケジュール」として提示することとした (図 7)。過去のライフログから時間帯毎にユーザが行なっている確率の高い行動を最大 3 つまで抽出し、各行動に対して最も値の大きい介入を選択してメッセージを作成する。例えば図 6 の方策が出力された場合、ユーザが 23 時台に行なっている確率の高い行動として風呂・運動・休息の 3 つがあるとき、エージェントの介入としては風呂・休息の場合は運動をするよう促し、運動中の場合は介入しないことが最適な戦略となる。他の時間帯についても同様に処理を行い、ユーザには図 7 のようなメッセージに整形して slack で通知する。図 7 では 23 時以降に関するアドバイスはあるが、それより前の時間帯については何も記載されていない。これは、ユーザがどんな行動を行っていてもエージェントは「介入なし」が最善の選択となる時間帯があり、その時間帯はメッセージ内容か



図 6: 方策の例. 時間帯 (0~23 時) × ユーザが行なっている行動 × エージェントの介入から構成される.

表 2: 実験スケジュール

日付	内容
6月8~9日	説明会・事前アンケート実施
6月9~12日	試験運用期間
6月15日~	実験開始
6月29日~	介入有無の切り替え
7月13日	実験終了
7月14日~	終了後アンケート, インタビュー実施

表 3: 実験期間と収集データとの対応関係

期間	収集データ
開始前	事前アンケート
試験運用期間	ライフログデータ
介入あり期間	ライフログデータ, 日次アンケート (3 問)
介入なし期間	ライフログデータ, 日次アンケート (1 問)
終了後	アンケート, インタビュー

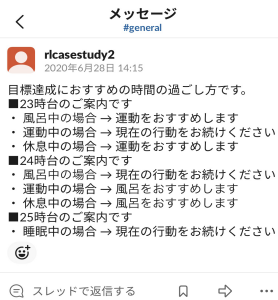


図 7: 生成されたおすすめスケジュール例. 例えば 22 時台にお風呂に入っていれば運動に切り替えるように, 運動中であればそのまま継続するように促している.

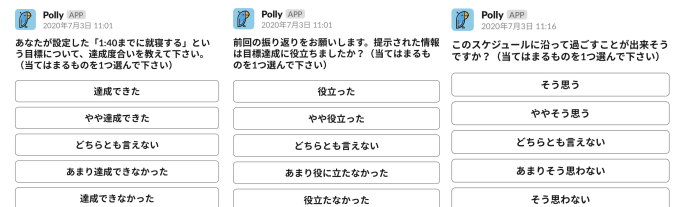


図 8: 目標達成度 図 9: 有用性 図 10: 実行可能性

ら除くように実装しているためである. 方策は実験開始から前日までのユーザのライフログを使って 1 日 1 回算出しており, 同じユーザであっても日によって生活パターンが変わると通知内容に多少変化が生じることがある.

4. 実験設計

4.1 実験の流れ

本実験の目的は, 我々のシステムを利用することで目標達成にどう影響するかを明らかにすること, また, 今後システムを洗練させるための知見を得ることである. 実験全体のスケジュールを表 2 に示す. 参加者は研究所内で 10 名募集し, 20~30 代の男性 6 名女性 4 名である. あらかじめランダムに半数ずつ 2 つのグループに分けた. 片方は介入有無の切り替え日までの最初 2 週間介入あり, その後 2 週間は介入なし. もう片方のグループは先に介入なし期間, 後で介入あり期間として実験を行なった. 介入あり期間中は朝夕におすすめスケジュール提示が行われるが, 介入無し期間中は提示は行われない. 試験運用期間は初期設定の確認とライフログを記録する作業に慣れるために設けており, 入力などに関する疑問をその間に解消すると同時に, 先に介入あり期間が来るグループのおすすめスケジュールを生成する初期データとして利用した. 実験前に事前アンケート, 実験期間中に 1 日 1~3 問のアンケート, 実験終了後にはアンケートとインタビューを実施しており, 今回の初期評価はこれらのデータを用いて分析した.

4.2 収集データ

参加者から収集するデータは, 毎日のライフログ, 実験前アンケート, 実験期間中アンケート, 実験終了後アンケートおよびインタビュー回答である. 実験の各フェーズと収集するデータとの対応関係は表 3 の通りである.

実験期間中に取り組む就寝時刻の目標は事前アンケートにて設定し, 実験中は毎日同じ目標に取り組むこととした. システム上はどの行動でも目標に設定することが出来るが, 今回就寝時間を題材とした理由は, 厚生労働省が推進している 21 世紀における国民健康づくり運動 (健康日本 21) [1] でも休養に関する項目が取り上げられ, 「十分な睡眠による休養の確保」が目標になっていることや, 世界的に見て日本人の睡眠時間が短いことから支援の必要があると考えたためである [8].

介入あり期間の日次アンケートは, 目標達成度合い (図 8), 提示されたおすすめスケジュールの有用性 (図 9) および実行可能性 (図 10) に関するものである. 介入なし期間の日次アンケートは目標達成度合い (図 8) のみ実施した. 日次アンケートはおすすめスケジュールと同様に slack で配信し, 今回の実験では Polly *7 を利用してアンケートを作成した.

終了後のインタビューでは, 介入あり期間となし期間とで目標達成や生活習慣に関する行動・意識の変化は生じた

*7 <https://www.polly.ai/>

か、このシステムにどのような改善が行われたらどんな場面で利用したいか、などについて今後の方向性を検討するために調査した。事後アンケートでもインタビュー内容と類似の質問を行っており、その回答をベースとして、収集データでは読み取ることが難しい意識や行動の変化を引き出すように努めた。インタビューの内容は全て録音し、それを書き起こしていくつかのカテゴリに分類して分析した。

4.3 評価指標

おすすめスケジュール提示の有無が目標達成度合いにどう影響するかについては、実験期間中の目標達成度合いに関する日次アンケート(図8)を利用して分析した。具体的には「あなたが設定した『～時～分(各自設定した目標時刻)までに就寝する』という目標について、達成度合いを教えてください。」という設問で、選択肢は(1)達成できた、(2)やや達成できた、(3)どちらとも言えない、(4)あまり達成出来なかった、(5)達成出来なかった、の5段階である。この回答は毎日集計しており、介入あり期間となし期間で行動の変化を比較した。

また、目標達成に対する意識の変化を調査するために、事後アンケートでおすすめスケジュール提示の有無が目標達成の難易度にどう影響しているかを尋ねた。具体的には、介入あり期間について「今回立てた就寝時間の目標を『おすすめの時間の過ごし方』があれば、達成するのはどのくらい難しいと感じますか?」、介入なし期間について「今回立てた就寝時間の目標を『自力で』達成するのはどのくらい難しいと感じますか?」という質問を行い比較を行った。回答の選択肢はいずれも(1)まったく難しい(10回やれば8回以上達成できると思う)、(2)やや難しい(10回やれば6回以上8回未満達成できると思う)、(3)普通である(10回やれば4回以上6回未満達成できると思う)、(4)やや難しい(10回やれば2回以上4回未満達成できると思う)、(5)とても難しい(10回やれば2回未満達成できると思う)、の5段階である。

5. 結果と考察

5.1 目標達成度合い

実験中に毎日収集した達成度合いに関するアンケートを用いて、介入あり期間と介入なし期間とで対応のあるt検定を行い達成度合いを比較したところ、10名中3名で、介入あり期間に有意に達成度合いが高まった(ID2: $t(8)=2.98, p<.05$; ID4: $t(9)=2.69, p<.05$; ID8: $t(9)=4.88, p<.01$, 5段階リッカート尺度:1=達成出来た, 5=達成出来なかった)。介入により達成度合いが変化した理由について、参加者はインタビューで次のように述べている。

[ID:2] ない時の方は意識しなかった気がする。来ると意識するので、意識するのは大事だなと思った。今家事した

方が良い、と書かれることが多くてぼーっとしてるとしなきゃなと思うし、実際家事やると早く寝れたので、今やっておくかという気持ちになった。

[ID:4] 介入期間中は目標時間に寝れることが多かったように思います。実際に寝れたという実感が増えたという感じがあります。夜に通知が飛んでくるようになって、仕事では忘れていたんですけど、もう1回改めて見て「お風呂の時間だ」などと、時間をより気にするようになりました。

[ID:8] 通知内容は自分の寝る前のルーティーンを全体的に前倒しするものだった。一番最初の提示内容はなるほどと思って、それは意識しようと思ったのは覚えています。

また、目標達成という行動の変化が生じたか否かに関わらず、システムがおすすめスケジュールを提示することにより、目標達成に対する難易度の感じ方が有意に下がったことが分かった(対応のあるt検定, $t(9)=2.69, p<.05$, 5段階リッカート尺度:1=難しくない, 5=とても難しい)。インタビューの中で、難易度が下がったと感じた理由について、参加者は次のように述べている。

[ID:1] 推薦されたおすすめの過ごし方を見て、実際にそれに沿える日もあって目標達成できていたので、なんだか結構達成できるし、これならできそうかなというのがありました。

[ID:6] 推薦システムがあればやりやすくなるかなと思った。寝る時間を意識するのに役立ったかもしれない。

[ID:9] ガイドがあること自体は嬉しい。ずっと覚えているわけではないけれど、頭の隅において行動したことはあった。ふと頭をよぎることがあった。

[ID:10] 基本的に22時に寝ることを意識していて、21時くらいから寝る準備をすると良いということは、提示される前から自分でもそう思っていた。今回実験で提示されて確信が高まった。通知をもらってから行動に確信が持てた。

このように、おすすめスケジュールの内容を実際に試して上手くいった経験や、納得感があつた場合は行動の変化に繋がりがやすく、それ以外でも出来そうという自信を得た場合や意識付け出来た場合には、目標達成に対する難しさを低減させる効果があつたものと考えられる。

5.2 システム改善の方向性

実験後のアンケートで、現在のシステムに関する評価を行った結果を図11に示す。各項目に対応する質問として、能力向上については「このシステムは自分の目標達成能力を高めてくれる」、有用性は「私の睡眠目標達成に役立つシステムだ」、重要性は「睡眠目標を達成するのに、このシステムを使うのは重要だ」、利用意向は「このシステムを利用できるなら、利用したいと思う」と尋ねた(7段階リッカート尺度:1=非常に当てはまらない, 7=非常に当てはまる)。これらの質問は技術の受容性に関するTAM(Technology

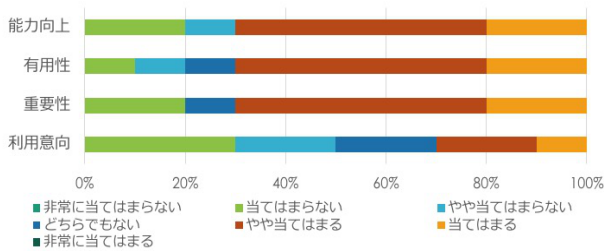


図 11: システムに関するアンケート評価結果.

Acceptance Model) [9] を参考にしている.

図 11 の結果を見ると、能力向上・有用性・重要性の 3 項目はほぼ同じ傾向であり、システムを利用することによって目標達成に効果があるのではないかと感じた参加者が多いと考えられる。一方で利用意向を見ると、現状のシステムのままではあまり利用したくないと感じた参加者が多い。事前アンケートで「就寝時間の目標を達成したい気持ちの強さについて、『絶対に目標を達成したいと思う』言葉が当てはまるか」について尋ねた際に、全く当てはまらないと答えた参加者はいなかったことから、そもそも目標達成する意欲が全く無いので利用したくない、ということはあるに (M=2.6, SD=0.8, 5 段階リッカート尺度, 1=よく当てはまる, 5=全く当てはまらない).

そこで、今後システムがどのように改善されれば利用したいと感じるかをインタビューで尋ねた。改善点として多かった意見は主に 2 つあり、目標設定の柔軟性と、ライフログの入力負荷に関するものである。

[ID:4] 入力のコストに対して、少し目標を変えると変わった提示が出るというのがありモチベーションになるし、達成しやすくなると思いました。これは明らかに無理だなというのが出てきたら目標を変えることもできますし。

[ID:5] 効果はすごく大きいと思うのですが、入力が一番最初が一番大変な感じでした。もうちょっと自動センシングしてくれて、ある程度埋まったところで正解不正解を自分で判断するくらいのシステムだともう少し楽な気がします。

[ID:7] 目標が色々設定できたり、入力が簡単で、ちゃんと効果がありそうだなという提示があればシステム自体の意味はあると思う。

インタビュー中に参加者に提示した入力負荷軽減策の中で好意的な反応だったものとしては、あらかじめライフログのテンプレートを作成しておきそれを修正する、fitbit などのデバイスで自動的に取得できる行動は自動で反映する、入力するとしても 1 週間分程度に留める、というものであった。今後は、今回の参加者のように既にライフログが蓄積されている場合にはそれをベースとしてモデル化し、ログが無い場合であっても 1 週間程度の入力に留め、日によって変わる予定など変更したい部分のみ入力する機能の追加を検討している。また、目標設定の柔軟さに関しては、アルゴリズム上はどの行動・時間でも目標に設定す

ることが可能なため、目標変更の入力を受け付ける機能の実装を検討している。今回は睡眠を題材としたが、インタビューでは運動習慣の形成に使いたいといった声もあり、行動・時間を変更できるようになればより利用が広がるものと考えている。

6. まとめ

本研究では、我々がこれまでに構築した強化学習により最適な介入方策を学習する技術を実装したシステムを用いてユーザ実験を行い、目標達成を支援する効果があるのか、また今後どのような改善をすれば良いかについて分析を行った。睡眠に関する生活リズム改善を題材とした実験から、おすすめスケジュール提示によって目標達成度合いが有意に高まった参加者が 10 名中 3 名おり、他の参加者についても目標達成に対する難しさを有意に低減させる効果があることが分かった。今後は、目標設定を柔軟に行えるようにすること、そしてライフログ入力負荷の軽減によりシステムの利用意向が高まるものと考えており、システムを改良した後に大規模に効果検証を行うことを検討している。

参考文献

- [1] 厚生労働省 生活習慣病予防, 入手先 https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryuu/kenkou/seikatsu/seikatusyuukan.html (2020.08.07).
- [2] Sutton, Richard S., and Andrew G. Barto. Reinforcement learning: An introduction. MIT press, 2018.
- [3] 高橋公海, 幸島匡宏, 倉島健, 戸田浩之. 強化学習を用いた健康行動促進のための介入戦略の学習. DEIM Forum. 2020.
- [4] Daskalova, N., Lee, B., Huang, J., Ni, C., and Lundin, J. Investigating the effectiveness of cohort-based sleep recommendations. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2018.
- [5] Takahashi, M., Kohjima, M., Kurashima, T., Matsubayashi, T., and Toda, H. Identifying self-changeable actions toward regulating rhythm of daily life. In Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers (pp. 218-221).
- [6] Puterman, Martin L. Markov Decision Processes. Discrete Stochastic Dynamic Programming. John Wiley and Sons, 2014.
- [7] Haarnoja, T., Tang, H., Abbeel, P., and Levine, S. Reinforcement learning with deep energy-based policies. In Proceedings of the 34th International Conference on Machine Learning. JMLR. org. 2017.
- [8] Organisation for Economic Co-operation and Development (OECD). National Time Use Surveys, 入手先 <http://www.oecd.org/gender/data/balancingpaidworkunpaidworkandleisure.htm> (2020.08.07).
- [9] Venkatesh, Viswanath, and Fred D. Davis. A theoretical extension of the technology acceptance model: Four longitudinal field studies. Management science 46.2 (2000): 186-204.