（資料１）

消費者とAI
－報酬環境と社会的側面－

唐沢かおり
東京大学名誉教授
工学系研究科上席研究員

本日の話題

事務局からのご依頼事項

「報酬」や「強化学習」に関する心理学の基本

本専門調査会で上がった消費者問題の事例の背後にある心理的過程について

（可能であれば）消費者の自衛策の一つとしての認知コントロール

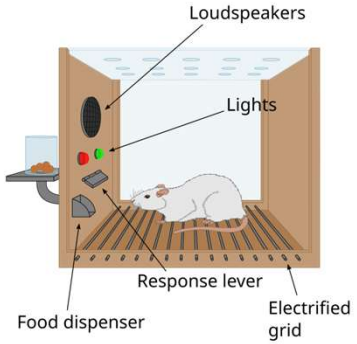
本日の話題

報酬や強化学習に関する基本の整理

社会的な側面に着目した「消費者×AI」をめぐる心的過程

行動主義と強化学習研究

- 報酬・強化：20世紀初頭以降の学習心理学(行動主義心理学)における中心的な研究テーマ
- ソーンダイクの「効果の法則」(Thorndike, 1898)
  - 仕掛け箱に入れた猫が、試行錯誤の末にレバーを引くなどして脱出する様子を観察
  - 満足をもたらす結果（脱出→餌）を伴った行動は、その後くり返されやすくなる
- スキナーの「オペラント条件づけ」(Skinner, 1938)
  - レバー押しなどの自発的行動と、その結果(餌など)の随伴関係を体系的に検証
  - 「結果(強化子)が伴う行動は増え、伴わない行動は減る」ことを実験的に示した
  - 弁別刺激（緑ランプの時のみレバー押しで餌）→特定のサイン下の行動のみ効果



- 「行動は意図や意思だけでなく、その結果との関係によって形づくられる」という理解の基盤となった
- 報酬・強化という視点は、人がなぜある行動を獲得するのか、繰り返すのかについて、意図や意思ではなく、行動とそれがもたらす結果との関係から説明するもの

心理学における強化学習の基本メカニズム

学習を進める3要因

- 強化子
  - 行動の後に生じ、その行動を増加、減少させる結果
- 予測誤差
  - 実際の結果 – 予測していた結果
  - 予想以上によい結果ほど、その行動の価値が上がる
- 強化スケジュール
  - 強化子が「いつ・どの頻度で」与えられるか

基本過程

- 行動の実行
  - 探索的に実行
  - すでに価値が既知の行動を実行
- 行動結果の経験、予測との差を評価
  - 欲求や期待との観点から、快・不快・価値を経験、評価
- 行動価値の更新
  - 行動がもたらしうる結果（価値）の更新、高いと選択されやすくなる

強化学習がもたらす結果

- 選択の偏り・反復・持続
  - よい結果につながった行動が、次回も選ばれ、維持されやすくなる
- 習慣化・自動化
  - 反復が続くと、習慣化、手がかりに反応した自動的行動化
- 行動の固定化
  - 即時に強化される行動が、他の選択肢より優先されることがある

強化学習とは、行動の結果と予測との差をもとに、その行動を再び選ぶ価値を更新していく過程

報酬・罰・強化子とは

- 報酬・罰：一般的に、よいもの、ご褒美、不快なこと・・・それにより、行動の変容が起こると「考えられている」
  - 褒められた行動を繰り返す、叱られた行動をやめる・・・
- 強化子：学習心理学の文脈では、ある行動のあとに生じ、その行動の頻度が増える（正の強化子）、または減る（負の強化子）という事実により定義される刺激のこと（効果を持った報酬や罰）
  - 「一次性強化子」（無条件性強化子）：食物、水、痛み刺激のように生得的に価値をもつ
  - 「二次性強化子」（条件性強化子）：金銭・称賛・非難・情報取得のように経験を通じて価値を獲得したもの
  - →「身体的な快・不快」の増減だけではなく、安心感、達成感、承認、拒絶、注意を向けられること、情報が得られること、不安や退屈の増減も、行動を強める場合がある
- 強化子とは人の行動を方向づけ、くり返しやすくする・頻度を減らす「結果」・・・行動の変化から定義される概念、
  - 何が人の行動を変える強化子になるかは、個人の状態、文脈、経験によって変わる

正の強化・負の強化・正の弱化・負の弱化（罰）

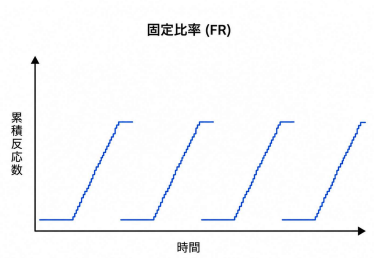
- 正の強化：行動のあとに望ましい刺激や結果が加わり、その行動が増える
  - 役立つ情報が得られる、称賛される、ポイントが得られる、楽しい体験が得られる
- 負の強化：行動によって不快な状態が取り除かれ、その行動が増える
  - 不安が和らぐ、退屈がまぎれる、迷いが減る、面倒な比較をしなくてすむ
  - 「負の強化」は罰ではなく、不快な状態の除去によって行動が強まることを指す
- 正の弱化：行動の後に不快な刺激が加わり、その行動が減る
  - ルール違反をしたら罰金を取られたので、以降違反をしない
- 負の弱化：行動の後に望ましい刺激や状態がとりのぞかれ、その行動が減る
  - 遅刻を繰り返したらボーナスが減ったので、以降、遅刻しなくなった

どのようなスケジュールで与えられるかが行動頻度・パタンを決める

強化スケジュールの基本 (Ferster ＆Skinner, 1957)

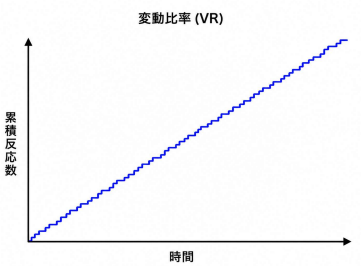
固定比率 Fixed Ratio(FR)

- 一定回数の反応ごとに強化
- 例：5回買うと一つ無料
- 反応率は高いが強化直後に休止が生じやすい・目標勾配が働く（近づくとがんばる）



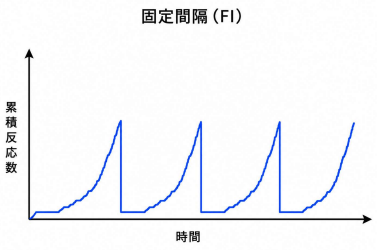
変動比率 Variable Ratio (VR)

- 平均すると一定回数ごとだが必要反応数は毎回変動
- 例：スロットマシン
- もっとも高く安定した反応



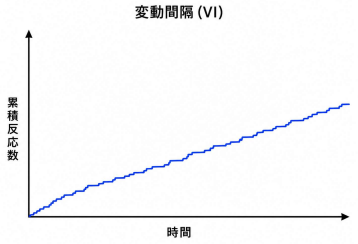
固定間隔 Fixed Interval(FI)

- 一定時間経過後の最初の反応を強化
- 例：タイムセール
- 時間が近づくと反応増加、行動組み込みが行われやすい



変動間隔 Variable Interval(VI)

- 平均すると一定時間ごとだが強化時点は毎回変動
- 例：不定期に届くクーポン
- 中程度で安定した反応が続く、こまめなチェック、習慣化



AIと強化環境

- 従来から、強化子（報酬）や強化スケジュールは消費者に対して用いられてきたが・・・
- AIにより変わること
  - AI自体が個人の反応を観察し、何をどのタイミングでどのように与えれば、行動変容につながるかを学習できる
    - 一律の「報酬」が個別最適化された「強化子」になる
    - 報酬提示による行動形成・変容過程がAIにより「量的に加速」
  - 消費者と会話（相互作用）しながら消費行動に影響する働きかけを行う
    - 会話から消費者の心的状態（関心、感情、価値観など）を推定
    - それに応じて説得的メッセージを個別生成し応答に応じて調整
    - 個別最適化された「相談・対話」も消費の対象となる（社会的欲求を満たすなど）
    - 選択の意味や理由そのものを構成するという「質的変化」・・・AIが選好そのものを形成する主体化
      - ←AIが選好の形成を「助ける」と考えれば、ネガティブだとは言い切れない

相談の消費：負の強化と回避学習の観点から

- 負の強化：行動によって不快な状態が取り除かれ、その行動が増える
- 回避学習：嫌なことが起こる前に、それを予測して回避する行動を獲得する
  - Mower（1947）の恐怖条件付けと回避学習
    - 不安が特定の状況と結びつくと（音が出ると電気ショック）、それを低減するための回避行動が負の強化で維持される（音が出るとショックを受ける前に逃げる）

- 商品を選択し購買する場面での後悔
  - よく調べすに購入して失敗した経験→購入場面自体が不安を喚起
  - 不安を感じた際にAIに「相談」し、安心を得ると、不安の軽減自体が（商品自体ではなく）相談行動を強化
  - これが定着すると、購入場面が「脅威場面化」・・・事前に不安を「予期」し、それを回避する行動として「AIに相談する」ことを学習する
- 消費者行動が、モノやサービスの購入としてだけではなく、「相談する」「確認する」という行為としてもとらえられる
- 安心するための「相談」行為自体が、繰り返し消費される対象となる

AIへの相談は、なぜ問題をもたらしうるのか

- 通常の対人相談には、負の強化のループを抑える自然な歯止めが（相対的に）かかりやすい
  - 何度も同じことを聞くと気まずい、相手の負担になる、相手が対応できる時間には限りがある、といった社会的コスト
- AIへの相談には、こうした歯止めがなく、いつでも、何度でも、疲れも見せずに応答が返ってくる
- →購買場面自体が脅威であるという信念から逃れられない
  - 不安を下げるための行動の過度な継続が、長期的には不安を維持してしまうパラドックス（Salkovskis, 1991）
  - 脅威を感じる場面で、「安全行動」により一時的に不安が下がると、安全であった理由を行動に帰属し、その場面自体が脅威をもたらすものだという信念は維持される
  - 安全行動（AIへの相談）をとり続けると、「実際には相談しなくても大丈夫」と学習する機会が失われる
- AIへの反復相談が、歯止めなく繰り返される「安全行動」として機能し、不安を自己強化的に維持・悪化させる回路を生み出す可能性
  - AIへの相談が、便利な情報収集から不安の維持システムに変容

AI自体を消費する行動が起こりやすい

- 脅威や不安がなくても、なぜAIを使ってしまうのか？
- 行動を生じさせる条件がそろいやすい（十分な動機づけ、その行動を実行できる容易さ・能力、適切なきっかけが同時にそろいやすい (Fogg, 2009）
  - 動機付け　　問題解決、不安低減、承認、好奇心、孤独の軽減
  - 容易さ　　低コスト、即時、24時間、対人負担が少ない
  - きっかけ　　通知、画面への組み込み、迷い・不安などの内的きっかけ
- 報酬が得られやすく、行動が習慣化しやすい：即時の答え、安心、効率化、時に予想外の有益な回答
- 生成AIで特に重要な点：相互作用を通じて動的に調整される可能性
  - 利用者の関心や不安を推定し、その人にとって意味のある答えを生成することで、「このAIは自分を理解してくれている、話を聞いてくれている」と感じる→社会的動機を満たすことで、再利用を高める可能性
  - 利用履歴や文脈を保持により、AIが先回りして候補を出すなど、利用者自身が考えたり選択肢を探索したりするコストを減らし行動を容易化する

社会的行動としての消費とAI

消費の社会的意味

- 自己の表現・構成
  - 何を選ぶかが自分らしさをつくる(Belk, 1988)
- 所属・同一視
  - どの集団に属し、どのような価値観を共有しているかを示す(Escalas & Bettman, 2005)
- 承認・地位
  - 他者からの評価、威信、自らの社会的地位を示す(Veblen, 1899)
- 規範・社会的証明
  - 多くの人が選ぶものを消費することで正しさを確認する (Cialdini, 1984)

「信頼できる」影響エージェントとしてのAI

- 社会的意味付けは、他者からの働きかけで、より強固な消費原理となる
  - 「あなたのために選びました」(好意・パーソナライズ)
  - 「同じような人が選んでいます」「多くの人に支持されています」(社会的証明)
  - 「この選択があなたらしい」(自己アイデンティティ)
- →社会的意味を満たす消費を促進する働きかけ
- AIが可能とする親身で、即時的で、個別的に見える応答は、信頼を生みやすく、信頼されるエージェントはより大きな影響力を持つ

認知コントロールと自衛策

- 認知コントロール：その場で目立つ刺激や即時的な欲求にそのまま反応せず、現在の目標に合わせて、注意・思考・行動を調整する…つまり立ち止まって考えること
  - コントロールの構成要素とできそうなこと
    - 自動的・衝動的な反応をいったん止める…判断までの時間を一定置くことをルール化
    - 自分の目標や基準を意識する…予算、何のために買うのか、優先事項を明示する（書き出すなど）
    - 別の見方や選択肢に切り替える…他の選択肢や買わない理由（ダメなところ）を意識的に探す
  - AIが作った選択環境から一歩離れ、自分の目標と基準に基づいて再評価できる状態を確保する

- ただし、AIによる働きかけが継続的で、個人化され、見えにくい現状を踏まえると、個人の注意力や抑制力で抵抗できると考えるのは危険
- これまで議論されてきたこと（インターフェース設計、透明化、パーソナライズの拒否、脆弱性を利用したターゲティングの制限など）が重要

文献リスト

- Belk, R. W. (1988). Possessions and the Extended Self. *Journal of Consumer Research*, *15*(2), 139–168.
- Cialdini, R. B. (1984). *Influence: The Psychology of Persuasion*. William Morrow
- Escalas, J. E., & Bettman, J. R. (2005). Self-construal, reference groups, and brand meaning. *Journal of Consumer Research, 32*(3), 378–389.
- Fogg, B. J. (2009). A behavior model for persuasive design. *Proceedings of the 4th International Conference on Persuasive Technology*. 40, 1-7.
- Mowrer, O. H. (1947). On the dual nature of learning—a re-interpretation of "conditioning" and "problem-solving." *Harvard Educational Review*, 17, 102–148.
- Skinner, B. F. (1938). *The Behavior of Organisms: An Experimental Analysis*. New York: Appleton-Century.
- Ferster, C. B., & Skinner, B. F. (1957). *Schedules of Reinforcement*. New York: Appleton-Century-Crofts.
- Thorndike, E. L. (1898). *Animal Intelligence: An Experimental Study of the Associative Processes in Animals*. Psychological Review Monograph Supplements, 2(4), 1–109.
- Veblen, T. (1899). *The Theory of the Leisure Class*. Macmillan