題目：規範心理再訪：強化学習モデルによる探索的検討

氏名：笹川陽奈子

指導教員：竹澤正哲

現実社会において、多くの人は規範に従った行動（以下、規範行動）をしている。規範は社会の秩序を保ったり、大規模な協力を実現させたりする上で重要な働きをしているといえる。人の規範行動がなぜ生まれるのかという問いについては、従来さまざまな理論が提唱されてきた。たとえば、人は規範を内面化することによって規範に従うようになるとする理論や、社会学習や個人学習の結果規範に従うようになるとする理論などである。しかし、近年、人の規範行動は強化学習で説明できるとする主張が注目されるようになってきた。本研究ではこの視点に基づき、強化学習モデルから人の規範行動が生成され得るか、生成され得るのであればどのようなメカニズムが働いているのかをシミュレーションによって検討した。現実において、規範行動のほうが行動の期待値が低い場合であっても多くの人は規範に従っている。この状況についてくわしく調べるために、本研究では規範行動を「期待値は低いが確実な行動」と定義した。「確実な行動」としたのは、規範行動をとると、他者から褒められるなど何らかの形で報酬が得られると考えられるためである。シミュレーションは、課題の報酬と遷移確率をさまざまに変えて行った。まず、2 つの学習アルゴリズム（モデルベース学習 1、モデルフリー学習）についてそれぞれ検討したところ、規範行動の方が期待値が低い場合であっても、規範行動の学習がみられた。そのうち、モデルベース学習 1 は規範逸脱行動をしたときにそれが露見する確率を高く見積もっていることが、規範行動を生み出している原因であるとわかった。そこで、遷移確率の学習プロセスを操作したモデルベース学習 2 についてシミュレーションを行ったところ、規範行動は学習されにくくなった。また、モデルフリー学習ではリスクのある選択肢を避ける傾向から規範行動が生み出されたことがわかった。ほかにも、さまざまな条件でシミュレーションを行ったが、ほとんどの条件において同様の結果となり、結果の頑健性が示された。本研究の結果をもとに、シミュレーションの課題構造をもっと複雑にしたり、従来のモデルを強化学習の文脈で表現したり、現実の人間を対象とした実験を行ったりすることによって、人間の規範行動・規範心理を解明していくことができると考える。