

農作物の自動灌水制御に向けた ニューラルネットワークを用いたQ学習

Q-learning with Neural Network for Automatic Irrigation Control of Crops

難波 脩人 辻 順平 能登 正人
Shuto Namba Junpei Tsuji Masato Noto

神奈川大学大学院工学研究科電気電子情報工学専攻
Graduate School of Electrical, Electronics and Information Engineering, Kanagawa University

There are many studies to reproduce cultivation methods based on experience and intuition of skilled farmers by supervised learning. Farmer's knowledge has an impact on the collecting data to conduct supervised learning, and it is difficult to judge whether such knowledge is a true optimal solution in agricultural work. So, we focused on reinforcement learning (RL) which learns based on reward given from the environment as a learning method independent the knowledge like experience and intuition. There are few cases of study using RL, and also the effectiveness has not been enough clarified. In this paper, we tried to cultivate Komatsuna based on simple Q-learning as a preparation to apply RL to the plant cultivation. As a result of Q-learning for a Komatsuna, it was possible to confirm the state of giving water according to the height of plant.

1. はじめに

日本の農業が抱える問題の一つに農業従業者の高齢化による離農率の上昇があげられる。それに伴い、熟練農家の長い年月をかけて培った技術が継承されことなく失われていくことが懸念されており、熟練農家の知見を残すためにも技術継承は農業分野における課題である。こうした問題を解決するため、センサやIoT技術を農業に生かしたスマート農業が注目されている。これらの分野では機械学習を用いて従来経験と勘によって行われてきた農作業の方法を再現する取り組みが行われている。特に機械学習の中でも、センサやカメラから収集した温湿度や水分量、草姿の変化などのデータを用いて教師あり学習によって熟練農家の知見を再現しようとする研究が盛んに行われている。例えば、センサを用いた研究では茎の太さの変化を測定することで植物に付加されているストレスを推定する研究 [Gallardo 06]、露地栽培における土壌水分量や水耕栽培における養液の濃度を測定することで自動灌水制御を行う研究 [Abidin 14] などがある。一方でカメラを用いた研究では植物の画像から病害診断を行う研究 [Mohanty 16]、葉のしおれ具合を推定することからストレス推定を行う研究 [Kaneda 17] などがある。農業のような複雑な要因が絡む作業において、農家の行う栽培方法が最適かどうか判断することは難しい。一方で、センサやカメラを用いてデータを収集する際に植物のどこに注目し、どのデータが重要であるか判断するために熟練農家の知見が必要不可欠である。さらに、実際の栽培環境から教師あり学習に必要なデータを採取するには限界があることが知られている。そのため、熟練農家の知見は教師データの質に大きく影響し、採取したデータに基づいて学習した結果が最適かどうかの判断は困難である。

そのような背景のもと、熟練農家の知見である経験と勘に依存せずに環境から与えられる報酬に基づいて学習を実行する学習手法として強化学習があげられる。実際に植物栽培に強化学習を適用した研究としてシミュレーションによる収穫量の最適化 [Sun 17] や栽培を繰り返すことによる養液供給の最適

化 [Wakahara 10]などを目的としたものがある。実際の植物栽培に強化学習を適用した研究に焦点を当てた場合、栽培期間の長さや状態の時間変化など植物のもつ特徴によってロボティクスで用いられるような手法をそのまま適用することが困難であることが知られている。Wakaharaらの研究は植物の栽培に強化学習を適用した際に生じるこれらの問題を植物の状態遷移や報酬の与え方から定義することで解消している。しかしながら、その事例は少なく、有効性は十分明らかにされていない。

本研究では、植物の栽培に強化学習を適用するための準備としてQ学習に基づく小松菜の栽培を実践する。植物の状態は時間変化することから同じ状態は存在せず、全ての状態に対してQテーブルを作成することが困難である。そのため、ニューラルネットワーク (NN) を用いたQ関数の作成手法を採用する。植物栽培におけるNNを用いたQ学習の有効性や課題を検討する。

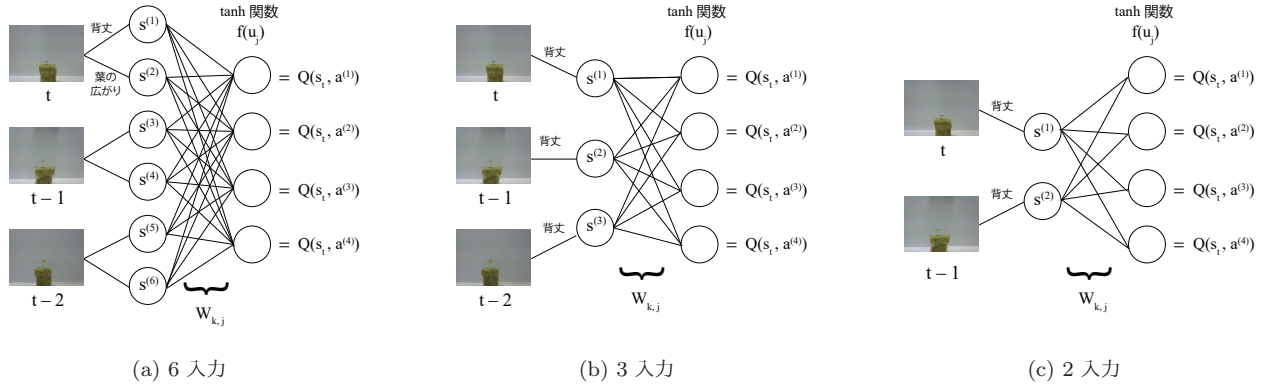
2. 植物栽培の最適化に向けた強化学習

2.1 ニューラルネットワークを用いたQ学習

本研究では植物がよりよく成長するための灌水行動の最適化に向けて強化学習の手法の中でも最もよく利用されるQ学習を用いる。Q学習は状態と行動の各組に対してQ関数を探索と活用を繰り返すことで更新し、エージェントが最適な行動を選択するよう学習させる手法である。

一方、植物栽培の最適化を強化学習によって行うにあたり重要な点として状態の定義があげられる。植物の状態は時間とともに変化することから同じ状態は1試行の中に1度しか現れない。つまり、Q学習ではすべての状態に対してQテーブルを構築することが現実的ではない。さらに、植物の状態は連続量であることから状態をQ関数で表すことが困難である。このような場合において重み W_{kj} を用いてQ関数の作成に関数近似を用いる手法が知られている。本研究におけるQ関数の導出概要を図1に示す。関数近似には出力層にtanh関数を用いたNNを使用している。複雑なNNで学習できるほど十分なデータが集められるとは限らないため、NNの構造は単層構造とした。Q関数はtanh関数を用いて近似するため、

連絡先: 難波脩人, 神奈川大学大学院工学研究科電気電子情報工学専攻 能登研究室, 〒221-8686 神奈川県横浜市神奈川区六角橋3-27-1, 電話: 045-481-5661

図 1: 関数近似を用いた Q 関数の作成概要

$f(u_j) = \tanh(u_j)$ としたとき、以下の関係式が成り立つ。

$$Q(s, a^{(j)}) = f(u_j) \quad (1)$$

$$u_j = \sum_{k=1}^n W_{kj} s^{(k)} + b_j \quad (2)$$

ここで時刻 t の状態 S_t の入力には 3 パターン用意し、それぞれ $n = 6$ 入力, 3 入力, 2 入力としている。植物の状態を定義する際に時刻 t における背丈や葉の広がりのみから状態を定義した場合、現在の状態に至るまでの遷移が反映されないことが考えられる。例えば、時刻 $t-1$ で健康な状態と比較して枯れているとみられる植物がある行動によって時刻 t で回復した場合と時刻 $t-1$ で健康な植物が同じ行動で健康なままの場合を考える。時刻 t の植物の状態のみを反映した s_t と時刻 t と $t-1$ の状態を反映した s_t とではより詳細な植物の状態を表すためには、後者の状態のほうが望ましいと考えられる。また、植物の状態は背丈だけに現れるのではなく、子葉の成長具合にも反映される。そのため、6 入力の場合は時刻 t , $t-1$, $t-2$ の各時刻における植物の背丈と葉の広がりを表した 2 次元ベクトルを 3 つ並べた 6 次元ベクトルとして $s_t = (s^{(1)}, s^{(2)}, s^{(3)}, s^{(4)}, s^{(5)}, s^{(6)})$ で表される。3 入力の場合は時刻 t , $t-1$, $t-2$ の各時刻における植物の背丈のみの 1 次元ベクトルを 3 つ並べた 3 次元ベクトルとし、2 入力の場合は時刻 t , $t-1$ における植物の背丈を 2 つ並べた 2 次元ベクトルとしている。学習時における重み W_{kj} の更新式は以下の式 (3) のようになる。また、式 (3) の δ_t の導出は式 (4) に示す。

$$W_{kj} \leftarrow W_{kj} + \alpha \delta_t \frac{\partial f(u_j)}{\partial W_{kj}} \quad (3)$$

$$\delta_t = r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t) \quad (4)$$

α は学習率, γ は割引率とし、それぞれ $\alpha = 0.5$, $\gamma = 0.9$ と設定している。報酬の定義はどの入力パターンに対しても同様に時刻 t と $t-1$ における植物の背丈を比較し、以下のように定義する。

$$r_t = \begin{cases} +0.2 & (\text{時刻 } t \text{ で成長している}) \\ 0 & (\text{変化なし}) \\ -0.2 & (\text{時刻 } t \text{ でしおれている}) \end{cases}$$

行動 (action) は $a^{(0)} = "0 \text{ ml}"$, $a^{(1)} = "10 \text{ ml}"$, $a^{(2)} = "15 \text{ ml}"$, $a^{(3)} = "30 \text{ ml}"$ の灌水を 4 パターン用意し、状態 s_t

で action を起こした場合に次の状態 s_{t+1} に遷移するのは水を与えてからの変化に時間がかかることから 2 日後と間隔を開けるよう設定している。農作物の栽培期間は 1 か月から数か月かかるものがほとんどであり、種を植えてから収穫までを 1 試行とした場合の学習効率は悪い。

本研究における 1 試行は学習効率を上げるために栽培期間を 2 週間として学習している。強化学習の適用先は小松菜とし、状態は種を植えてから一律 30 ml の灌水を行った 4 日後を s_1 と定義する。

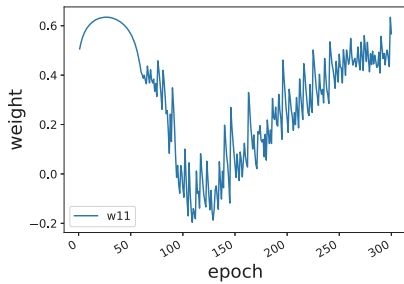
2.2 experience replay を用いた学習

1 試行やステップ数は学習の効率を上げるために 2 週間の 7 ステップと定義している。一方、シミュレーションに強化学習を適用した場合などと比較すると十分なデータ数が集まらず、植物に対する学習が最適解に収束することは困難であると考えられる。仮にステップ数を長くすることを考えると植物は成長段階によって要求する水分量が変化することから行動によって与える水分量を変化させる必要がある。強化学習を植物に適用していくためには植物の成長段階に合わせた学習手法が必要であると考えられる。本研究では植物が十分なデータを確保しつつ学習が収束していくことを目的とし、experience replay (ER)[Lin 92] を採用する。ER はエージェントが経験したサンプル $\{(s_t, a_t, r_t, s_{t+1})\}_{t=1}^T$ をメモリに保存する。このとき、 T はサンプル数の合計を表している。一定量のサンプルが保存されたところでメモリからサンプルをランダムに抽出し、学習の入力として再利用することで学習の効率化を図っている。つまり、メモリにサンプルを保存するために人の手によって小松菜の栽培を行うことでサンプルを一定量収集し、メモリに保存した栽培データを用いて重み W_{kj} を更新する。

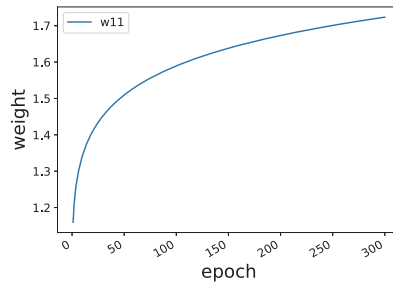
3. 実験および結果

ER を植物栽培に対して適用する際にメモリにどの程度のデータ数を保存し、かつデータをどの程度学習させることで植物が最適な方策を得るかを評価するため、3 つの栽培期間に分けてデータ採取を行った。小松菜は室内栽培用ハウスキットの中でロックウールと呼ばれる培養地に植えている。

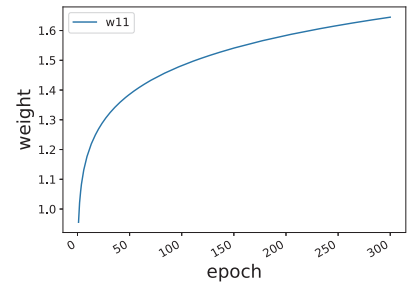
- 2018 年 12 月 1 日から 2018 年 12 月 13 日 (1 回目)
発芽しない個体や枯れる個体が現れることもあり、小松菜 48 株を栽培した結果、データ採取ができた個体は 28 株であった。このときのサンプル数の合計は $T = 168$ とした。



(a) 6 入力時の重みの更新の様子



(b) 3 入力時の重みの更新の様子



(c) 2 入力時の重みの更新の様子

図 2: ER による各入力に対する重み W_{kj} の更新の様子 ($k = 1, j = 1$ の場合)

- 2018 年 12 月 14 日から 2018 年 12 月 26 日 (2 回目)
小松菜 16 株栽培した結果, 15 株からデータを採取し, サンプル数の合計は 1 回目の結果と合わせて $T = 256$ とした.
- 2019 年 1 月 14 日から 2019 年 1 月 26 日 (3 回目)
小松菜 48 株栽培した結果, 40 株からデータを採取し, サンプル数の合計は 1 回目, 2 回目と合わせて $T = 496$ とした.

3.1 サンプル数 $T = 168$ に対する学習結果

実験では 3 つの栽培期間で採取したデータの各サンプル数に対して 6 入力, 3 入力, 2 入力の入力を試すことで植物に強化学習を適用するために必要なデータ数及び, 状態の定義など ER を用いた場合にどの手法が最適であるかを明らかにする. まず, サンプル数 $T = 168$ に対する ER の学習成果を評価するため訓練データに使用したデータを入力し, 予測された action と実際の action を比較した. その結果, スペースの都合上結果は除くが, どの入力の場合に対しても偏った行動しか示せず, 学習が成功した様子は確認できなかった.

3.2 サンプル数 $T = 256$ に対する学習結果

次に $T = 256$ のサンプルを訓練データとして ER を行い, 各入力に対する学習結果を評価するため, 同様に action の比較を行った. このとき, 3 入力と 2 入力の学習結果からは状態における action の選択にばらつきが生じていることが確認できた. 学習段階の様子を確認するため, 各入力に対する重み W_{kj} の更新の様子を図 2 に示す. 各グラフはそれぞれの重みに対して $k = 1, j = 1$ とした場合の値を表しており, 6 入力では重みが収束していない様子が確認できる. 一方で 3 入力と 2 入力の場合においてはどちらも収束している様子が明らかである.

3.3 サンプル数 $T = 496$ に対する学習結果

最後に $T = 496$ のサンプルを訓練データとして ER を行った結果に対して同様の評価を行った. 6 入力に対しては $T = 168, 256$ のときと同様にデータ数を増やしたところで改善は見られず, 学習回数を増やしたとしても重みが収束する様子は確認できなかった. 3 入力と 2 入力に対しては $T = 256$ と同様に重みの更新は収束している様子が確認できた. しかしながら, ER を行うにあたり, メモリに保存するデータは少ないほうが学習の効率は良いことが考えられるため, 本研究では $T = 256$ の場合に着目する.

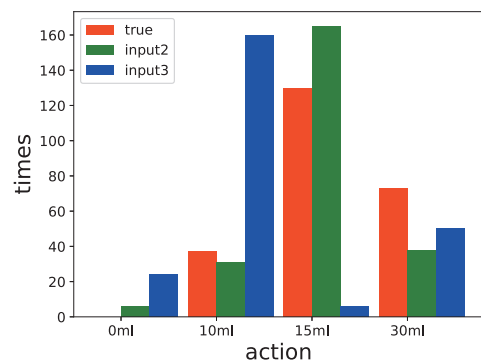


図 3: action の出現頻度

3.4 サンプル数 $T = 256$ かつ 2 入力に対する学習結果

3 入力と 2 入力における状態の定義としてどちらが優れているかを評価するために, 訓練データで使用したデータを NN に入力した際に出力される action のばらつきをまとめてグラフに示す (図 3). グラフの true は人の手によって小松菜を栽培した際に与えていた action を示している. input2 や input3 は強化学習を用いて学習した NN に訓練データを入力した際に得られた action を表しており, 後ろの数字はそれぞれ状態の入力数を表している. 横軸は action の種類, 縦軸は各 action の出現頻度を表している. グラフからわかる通り, input3 の action は true の action と比較すると $a^{(1)} = "10 \text{ ml}"$ の頻度が多く, $a^{(2)} = "15 \text{ ml}"$ の頻度が少ないことが明らかである. また, input2 の action と true の action を比較すると $a^{(1)}$ や $a^{(2)}$ の頻度はほとんど一緒であることが明らかであった. このことから 3 入力と 2 入力の場合では学習は同様に収束しているにもかかわらず, 2 入力の方が人の手による栽培と近い行動を選択していることが明らかである.

2 入力に対する強化学習の評価としてどのような方策に基づいて action を決定しているかを明らかにするために状態の入力に対する行動の選択分布をグラフに示す. 図 4 は強化学習による行動選択, 図 5 は人の手による行動選択を表しており, 横軸は入力時の現在の状態を表し, 縦軸は次の状態を表している. 強化学習による学習結果から得られた action は greedy 戦略に基づいて選択した結果を表しており, 入力に対して単純に最適な action を選択した場合の結果のみを表している.

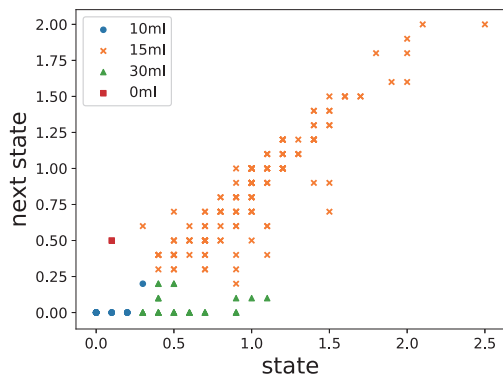


図 4: action の選択分布 (ER による結果)

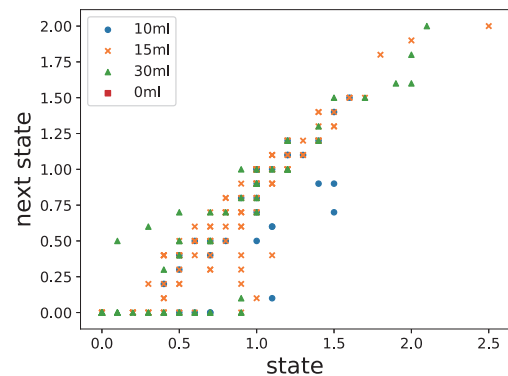


図 5: action の選択分布 (人の手による結果)

4. 考察

図 4, 5 から植物の状態を 2 入力とした場合の学習結果と人の手による栽培結果を比較すると、学習結果は人間と近い行動を選択している傾向がみられる。しかしながら、人の手による栽培では現在の状態と次の状態の組み合わせが同じ値をとっている場合においても異なる行動を選択している場合がある。原因として、人の手による栽培では植物の背丈や葉の広がり以外にもロックワールの乾き具合などを参考に与える水の量を決定しているが強化学習による Q 学習ではロックワールの乾き具合など、入力以外の外的要因を考慮していないことがあげられる。 Q 学習を植物栽培に適用した結果、図 4 のような分布がみられたことから植物は背丈の高さを基準に与える水の量を決めていることが明らかになった。この結果は一般的な Q 学習では困難であった植物の栽培に強化学習を適用する際の手法として NN を用いた手法が有効であることを示唆している。強化学習は環境に応じてエージェント自身が最適だと思う行動を選択することでその環境における最適化を求める手法であるため、今回のように人の手によって栽培することで得た行動を真値として比較することは必ずしも正当な評価方法とは言いきれない。しかしながら、図 3 に示したようにあまりにも人の手による栽培結果とかけ離れている場合は学習結果に妥当性はないと考えられる。人の手による栽培から得られたデータは現状の暫定解に過ぎず、強化学習によって学習した結果をもとに実際に小松菜の栽培を繰り返すことでエージェントが選択する行動の良し悪しを評価することが必要であると考えられる。

5. おわりに

本研究では農作物の自動灌水制御に向けて実際に小松菜の栽培に強化学習を適用することからその有効性や課題の検討を目的とした。植物の特徴である状態の時間変化や種を植えてから収穫までの栽培期間の長さを考慮するため、関数近似による Q 学習を採用して実験を行っている。さらに、栽培データの少なさをカバーしつつ、学習が収束していくことを狙いとして関数近似には NN の単層構造を採用し、ER による学習を行った。NN の入力には状態を 6 次元ベクトル、3 次元ベクトル、2 次元ベクトルと 3 パターン用意し、状態の定義として適切なものを明らかにした。実験を行った結果、状態の入力を 2 次元ベクトルとし、ER に用いるためのメモリの保存数 $T = 256$ とした場合の学習結果が人の手による栽培方法と最も近い行動を選択する様子が確認できた。よって、植物の栽培に NN を用

いた Q 学習を適用するためには状態の入力は 2 入力、メモリの保存数は $T = 256$ 程度とした場合、望ましいと考えられる action を選択することが明らかになった。しかしながら、人の手による栽培と比較したとき、強化学習による学習結果からは人間が判断材料としている複雑な要因を考慮している様子は見られず、改善の余地がある。

参考文献

- [Abidin 14] Abidin, M. S. B. Z., Shibusawa, S., Ohaba, M., Li, Q., and Bin Khalid, M.: Capillary Flow Responses in A Soil-plant System for Modified Subsurface Precision Irrigation, *Precision Agriculture*, Vol. 15, No. 1, pp. 17–30 (2014)
- [Gallardo 06] Gallardo, M., Thompson, R., Valdez, L., and Fernández, M.: Response of Stem Diameter Variations to Water Stress in Greenhouse-grown Vegetable Crops, *The Journal of Horticultural Science and Biotechnology*, Vol. 81, No. 3, pp. 483–495 (2006)
- [Kaneda 17] Kaneda, Y., Shibata, S., and Mineno, H.: Multi-modal Sliding Window-based Support Vector Regression for Predicting Plant Water Stress, *Knowledge-Based Systems*, Vol. 134, No. 15, pp. 135–148 (2017)
- [Lin 92] Lin, L.-J.: Self-improving Reactive Agents Based on Reinforcement Learning, Planning and Teaching, *Machine Learning*, Vol. 8, No. 3-4, pp. 293–321 (1992)
- [Mohanty 16] Mohanty, S. P., Hughes, D. P., and Salathé, M.: Using Deep Learning for Image-based Plant Disease Detection, *Frontiers in Plant Science*, Vol. 7, pp. 1–10 (2016)
- [Sun 17] Sun, L., Yang, Y., Hu, J., Porter, D., Marek, T., and Hillyer, C.: Reinforcement Learning Control for Water-Efficient Agricultural Irrigation, in *Proc. of ISPA/IUCC 2017*, pp. 1334–1341 (2017)
- [Wakahara 10] Wakahara, T. and Mikami, S.: Adaptive Nutrient Water Supply Control of Plant Factory System by Reinforcement Learning, in *Proc. of SCIS & ISIS 2010*, pp. 1020–1025 (2010)