修士論文

マルチエージェント強化学習を用いた複数ニューラル機械翻訳の自己組織化

指導教官　　村上 陽平 准教授

立命館大学情報理工学研究科
情報理工学専攻 計算機科学コース
6611210074-2

張 禹王

2022 年度 秋学期 特殊研究 4(CH)
令和 5 年 1 月 31 日

マルチエージェント強化学習を用いた
複数ニューラル機械翻訳の自己組織化

張 禹王

内容梗概

ニューラル機械翻訳の誕生により，機械翻訳の精度が向上し，世間から大きな注目を集めている．高精度のニューラル機械翻訳モデルを構築するには大規模かつ高品質の対訳データが不可欠である．ユーザ間での対訳コーパスの共有を必要としない連邦学習は，セキュリティや著作権などの理由で企業内に蓄積された非公開の大量の対訳コーパスの利用を促進すると期待される．しかしながら，連邦学習の参加者たちが保有しているデータは質やドメインなどが異なり非均質なため，一般的な連邦学習アルゴリズムのように全参加者のモデルを一括に統合しても大きな精度向上が期待できない．

そこで，本研究では連邦学習の集約プロセスのたびに，モデルの精度に応じて各参加者が動的に連携相手を選択しモデルを統合する自己組織化手法を提案する．具体的には，連邦学習の各参加者が，構築された参加者のモデルを集約するたびに，マルチエージェント強化学習によって獲得されたポリシーに従って，最適な連携相手を選択する．本手法の実現にあたり，取り込むべき課題は以下の 2 点である．

動的組織法

連邦学習の参加者が持つ学習データの特徴量と質に差がある．連携に適さない参加者と組織化してモデルを統合する場合，学習に参加したデータの総量が増えたとしても，構築されたニューラル機械翻訳モデルの精度が向上するとは限らない．精度を上げるために，集約するたびに特徴量と質が違う学習データを持つ参加者の存在を排除しながら，最適な相手と連携する自己組織化手法が必要である．

最善な組織パターンの探索

連邦学習に複数の参加者がいるため，参加者たちのモデルを組み合わせるパターンが多様である．ある組み合わせで得たニューラル機械翻訳モデルの精度が一時的に下がっても，最終的の精度が上昇することもありうる．また，連邦学習の参加者の増加と共に組み合わせの組織パターンも増大する．その結果，膨大な組み合わせパターンで最善な組織法を見つけ出すことが困難となる．将来

的なモデル精度の期待値を考慮し，組織化コストを抑える方法が必要である．

前者の課題に対しては，まず，連邦学習で集約するたびに，全参加者が組み合わせる組織パターンで組織して，パラメータサーバで集約を行い，複数の集約モデルを得る．そして，各参加者は自分の評価データをパラメータサーバに送って，各集約モデルの精度の評価を依頼する．最後にパラメータサーバは各参加者にとって精度が最も高い集約モデルを返し，参加者はそれぞれもらった集約モデルを使って次の学習に進む．

後者の課題に対しては，連邦学習の各参加者をエージェントとみなし，マルチエージェント深層強化学習を適用する．具体的には，集約されるニューラル機械翻訳モデルで評価データを翻訳し，評価データの各センテンスの BLEU スコアを状態空間とし，エージェントはどの参加者と連携するかをアクションとする．さらに，前回集約後のモデルから今回集約後のモデルの精度上昇量を報酬にする．このようにマルチエージェント深層強化学習を導入することで，更なる精度向上を図った．

提案手法によって構築されたニューラル機械翻訳モデルを，Federated Averaging で構築されたニューラル機械翻訳モデルと比較することで，提案手法によって構築されたモデルの精度増加率を用いて評価を行い，提案手法の有効性を検証した．本研究の貢献は以下の通りである．

動的組織法

動的組織法と Federated Averaging によって構築されたニューラル機械翻訳モデルを自ドメインと他ドメインの翻訳精度を評価し，動的組織法によって構築されたモデルが自ドメインの精度は 28.99%高くなり，他ドメインの精度にも 16.83%向上し，提案手法の有効性を検証した．

最善な組織パターンの探索

マルチエージェント深層強化学習の導入によって，連邦学習の参加者たちが将来的のモデルの精度上昇を考慮しながら，最適な連携相手を動的に選択することで自己組織化する手法を考案した．動的組織法によって構築されたニューラル機械翻訳モデルと比較し，6.33%精度を向上し，従来の中央集権型ニューラル機械翻訳の性能の 94.26%に達した．

Self-Organized of Multi-Neural Machine Translation Using Multi-Agent Reinforcement Learning

ZHANG Yuwang

Abstract

Large-scale and high-quality bilingual data are essential for generating highly accurate neural machine translation models. Federated Learning, which does not require the sharing of a bilingual corpus among users, will do so when it promotes the use of a large number of undisclosed bilingual corpuses accumulated within companies for reasons such as security and copyright. However, since the data held by Federated Learning participants are heterogeneous in quality, domain, etc. it is not expected to improve the accuracy significantly even if the models of all participants are integrated together like the general Federated Learning algorithm.

Therefore, in this study, we propose a self-organizing method in which each participant dynamically selects a partner for cooperation and integrates the model according to the accuracy of the model in each Federated Learning aggregation process. Specifically, each participant in Federated Learning selects the best partner to work with according to the policy acquired by multi-agent reinforcement learning, whenever the aggregate model be constructed. In realizing this method, the following two points should be incorporated.

Dynamic Organization Method

When models are integrated by organizing with participants who are not suitable for collaboration, the accuracy of the constructed neural machine translation model may not improve even if the total amount of data involved in learning increases. To improve accuracy, we need a self-organizing method that works with the optimal partner while eliminating the presence of participants with learning data whose feature quantity and quality differ each time they are aggregated.

search for the best organizational pattern

The patterns of combining the models of the participants in Federated Learning are diverse. A temporary decrease in the accuracy obtained with a certain combination may increase the final accuracy. When participants increases, the

best organization method with a huge combination pattern will be difficult to find. It is necessary to consider the expected value of future model accuracy and to find a way to reduce the cost of organization.

For the former problem, firstly, when each aggregate , all participants model be aggregate to multiple aggregation models.Then, evaluate the accuracy of each aggregated model. Finally, each participant proceeds to the next study using the aggrgated model with highest accuracy.

For the latter task, we apply multi-agent deep reinforcement learning. Specifically, the evaluation data are translated by an aggregated neural machine translation model, the BLEU score of each sentence of the evaluation data is made a state space, and the action is which participant the agent cooperates with. In addition, the amount of accuracy increase from the model after the previous aggregation to the model after the current aggregation is used as a reward.

By comparing the proposed method with Federated Averaging, the evaluation was carried out using the accuracy increase rate, and the effectiveness of the proposed method was verified. The contributions of this study are as follows.

Dynamic Organization Method

The neural machine translation model constructed by the dynamic organization method and Federated Averaging were used to evaluate the translation accuracy of the local domain and the other domain. The model constructed by dynamic organization method improved the accuracy of local domain by 28.99% and the accuracy of other domains by 16.83%, which verified the effectiveness of the proposed method.

search for the best organizational pattern

With the introduction of multi-agent deep reinforcement learning, we devised a method for Federated Learning participants to self-organize by dynamically selecting the best partner to work with while considering future model accuracy increases. Compared with a neural machine translation model constructed by the dynamic organization method, it improves the accuracy by 6.33%, reaching 94.26% of the performance of traditional centralized neural machine translation.

マルチエージェント強化学習を用いた
複数ニューラル機械翻訳の自己組織化

目次

第１章　はじめに

ニューラル機械翻訳（Neural Machine Translation，NMT）の誕生により，機械翻訳の精度が向上し，世間から大きな注目を集めている．高精度のニューラル機械翻訳モデルを構築するには大規模かつ高品質の対訳データが不可欠である．ユーザ間での対訳コーパスの共有を必要としない連邦学習 [1] は，セキュリティや著作権などの理由で企業内に蓄積された非公開の大量の対訳コーパスの利用を促進すると期待される．しかしながら，連邦学習の参加者たちが保有しているデータは質やドメインなどが異なり非均質なため，一般的な連邦学習アルゴリズムのように全参加者のモデルを一括に統合しても大きな精度向上が期待できない．

そこで，本研究では連邦学習の集約プロセスのたびに，モデルの精度に応じて各参加者が動的に連携相手を選択しモデルを統合する自己組織化手法を提案する．具体的には，連邦学習の各参加者が，構築された参加者のモデルを集約するたびに，マルチエージェント強化学習によって獲得されたポリシーに従って，最適な連携相手を選択する．本手法の実現にあたり，取り組むべき課題は以下の通りである

動的組織法

連邦学習の参加者が持つ学習データの特徴量と質に差がある．連携に適さない参加者と組織化してモデルを統合する場合，学習に参加したデータの総量が増えたとしても，構築されたニューラル機械翻訳モデルの精度が向上するとは限らない．精度を上げるために，集約するたびに特徴量と質が違う学習データを持つ参加者の存在を排除しながら，最適な相手と連携する自己組織化手法が必要である．

最善な組織パターンの探索

連邦学習に複数の参加者がいるため，参加者たちのモデルを組み合わせるパターンが多様である．ある組み合わせで得たニューラル機械翻訳モデルの精度が一時的に下がっても，最終的の精度が上昇することもありうる．また，連邦学習の参加者の増加と共に組み合わせの組織パターンも増大する．その結果，膨大な組み合わせパターンで最善な組織法を見つけ出すことが困難となる．将来的なモデル精度の期待値を考慮し，組織化コストを抑える方法が必要である．

以下，本論文では，２章においてニューラル機械翻訳と連邦学習について紹

介する．次に，３章では従来の連邦学習がニューラル機械翻訳に応用時に集約の問題について説明する．続いて，４章と５章はそれぞれ グリーディ法を用いたクライアントの動的組織手法と，マルチエージェント深層強化学習による自己組織化手法について具体的に説明する．

６章では，４章と５章で説明を行ったクライアントの動的自己組織手法に対する評価を行う．第７章はその評価結果について考察し，今後の展望や課題についてしらべて述べて結論する．最後に，第８章で本稿をまとめる．

第２章　ニューラル機械翻訳のための連邦学習

この章では，まずニューラル機械翻訳と連邦学習の概要について説明する．その後，連邦学習を用いたニューラル機械翻訳を紹介する．

2.1　ニューラル機械翻訳

ニューラル機械翻訳はニューラルネットワークを介して単語をある言語から別の言語に翻訳するアルゴリズムである．ニューラルネットワークは脳内の神経細胞（ニューロン）におけるネットワーク状の情報伝達の仕組みを模した計算モデルである．ニューラル機械翻訳は従来のルールベース機械翻訳（Rule Based Machine Translation，RMT）や統計的機械翻訳（Statistical Machine Translation，SMT）と比べ，翻訳精度を飛躍させた．

ニューラル機械翻訳は大きく分けるとエンコーダー（encoder），アテンション機構（attention mechanism），デコーダー（decoder）の３つの部分で構成されている．エンコーダーは翻訳される文書を入力文として順次に読み込んで，実数ベクトルに符号化する．そして，アテンション機構は出力単語を決める前に，符号化された入力文のどこに注目するかを調整する．最後のデコーダーは出力文を復号化する．図１はこれらの部分から構成される NMT の概略図である．[2]

ニューラル機械翻訳はモデルを学習する時に学習データとして対訳コーパスを使っている．その対訳コーパスが多ければ多いほど構築されたニューラル機械

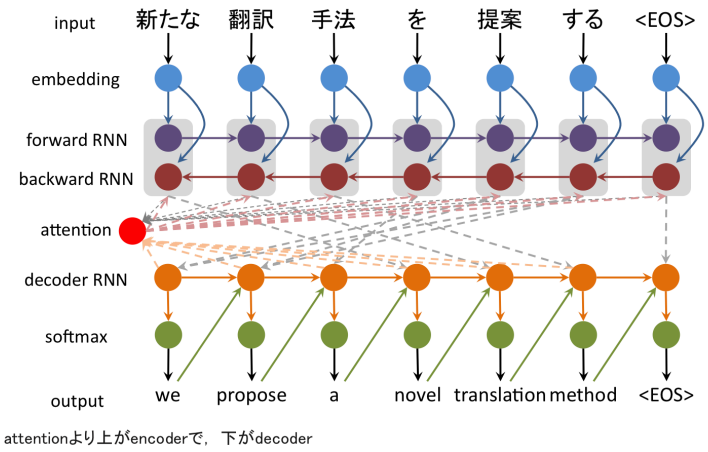

図 1: ニューラル機械翻訳の翻訳方式の概要 [2]

翻訳モデルの翻訳精度が高いという傾向があるため，ニューラル機械翻訳の品質が学習時に使う対訳データ量に強く依存している．しかし，現状ではセキュリティや著作権などの理由で大量の対訳データが企業内に蓄積され，一つの組織で十分な量の対訳データを入手することはとても困難である．連邦学習はニューラル機械翻訳において高品質な対訳コーパスが不十分という問題の一つの解決策として注目されている．

2.2 連邦学習

連邦学習（Federated Learning）は 2016 年にはじめて提唱された分散型の機械学習である．連邦学習に参加した各組織は自分が所有する学習データを共有せず，機械学習モデルの学習に必要な情報だけをローカルで計算し，共有する．連邦学習はこのことによりデータを一箇所に集中することなく，各学習データの特徴を反映できる機械学習モデルを生成する技術である．

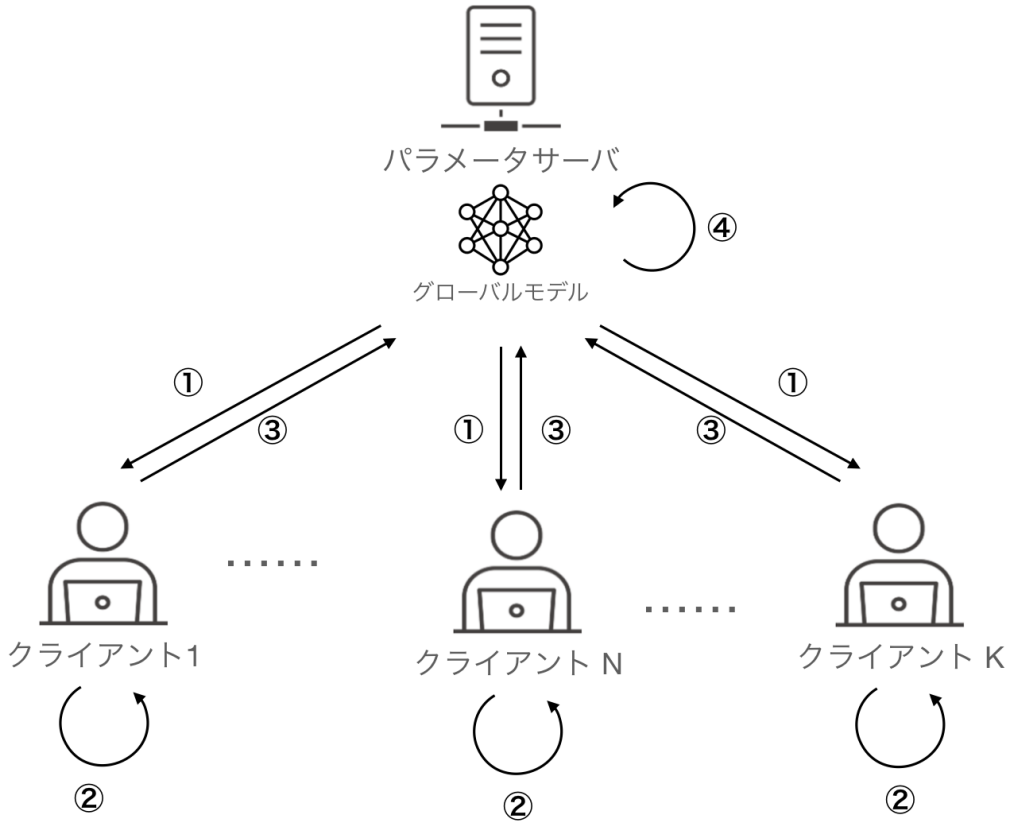

図 2: 連邦学習の流れ

一般的に連邦学習は複数のクライアントと一つのパラメータサーバから構成され，図 2 のように以下の流れで学習する．[1][3]

Algorithm 1 FederatedAveraging.

K はクライアント数; k はクライアントの番号; B はミニバッチサイズ; E は epochs 数; η は学習率; $\nabla l(\omega; b)$ はバッチ b についての損失 $l(\omega; b)$ の勾配

Server executes:

Initialize ω_0

for each round $t = 1, 2, \cdots$ **do**

$m \leftarrow \max(C \cdot K, 1)$

$S_t \leftarrow$ (random set of m clients)

for each client $k \in S_t$ **in parallel do**

$\omega_{t+1}^k \leftarrow \text{ClientUpdate}(k, \omega_t)$

end for

$\omega_{t+1} \leftarrow \sum_{k=1}^{K} \frac{n_k}{n} \omega_{t+1}^k$ ▷ モデルを集約

end for

ClientUpdate(k, ω)**:** ▷ client k で実行

$\beta \leftarrow$ (split P_k into batches of size B)

for each local epoch i from 1 to E **do**

for batch $b \in \beta$ **do**

$\omega \leftarrow \omega - \eta \nabla l(\omega; b)$

end for

return ω to server

end for

1. パラメータサーバは値がランダムに設定されている初期モデル（グローバルモデル）を生成し，各クライアントへ配布する．
2. 各クライアントは，受信した初期モデルに対して，自身が保有する学習データを用いてローカルで学習を行い，モデルを更新する．
3. 各クライアントは，更新されたモデルからモデルパラメータを抽出し，更新用データとしてサーバへ送信する．
4. サーバは，クライアントからもらった更新用データを集約し，グローバル

モデルのモデルパラメータを更新する．

5. サーバは，更新されたグローバルモデルを各クライアントへ配布し，2 の処理に戻る．

Federated Average (FedAvg) は連邦学習の代表的なアルゴリズムである．アルゴリズム 1 のように，サーバはランダムなクライアントを選択し，グローバルモデルを送り，学習させる．そして，クライアントからもらった更新用パラメータ （ω_{t+1}^{k} , k は該当クライアントの番号，t は学習回数）を各クライアントが持つサンプル数 n の重み付け平均によって統合している ($\omega_{t+1} \leftarrow \sum_{k=1}^{K} \frac{n_k}{n} \omega_{t+1}^{k}$)．

2.2.1 ニューラル機械翻訳への応用

連邦学習に参加した各組織は自分が管理している学習データを他人に共有する必要がない．そのため，プライバシーを保護し，セキュリティと著作権などの問題をクリアでき，複数の企業が共同でモデルを訓練することが可能になる．ニューラル機械翻訳において大規模かつ高品質な対訳コーパスが不十分という問題を解決できる．

実際「連邦学習を用いた非中央集権型のニューラル機械翻訳の有用性検証」で連邦学習で生成されたニューラル機械翻訳モデルはクライアントが連邦学習に参加せず，単体学習で生成されたニューラル機械翻訳モデルと比べ，クライアントたちの学習データのドメインが近い（均質データ）の時に精度が 8.94 % 上昇し，クライアントたちの学習データのドメインが遠い（非均質データ）の時にも 8.74 %向上することを分かる．ニューラル機械翻訳における連邦学習の有用性を検証できた．[4]

また，連邦学習をニューラル機械翻訳へ応用する際に計算上と通信の両方で非効率になる可能性があることに対して,「Communication-Efficient Federated Learning for Neural Machine Translation」ではニューラル機械学習の中にコントローラというレイヤを追加する手法を提案した．コントローラは連邦学習サーバと通信するための連絡係として役を立ち，連邦学習クライアントがモデルに対して学習する際に，最小限の更新情報を生成する．ニューラル機械翻訳における連邦学習のコストを削減できた．[5]

第3章　集約問題

先行研究でニューラル機械翻訳における連邦学習の有用性を証明した．しかしながら，実際に連邦学習の参加者たちが保有しているデータは，質，特徴量などが異なり非均質なため，一般的な連邦学習アルゴリズムのような全参加者のモデルを一括に統合する手法で構築されたニューラル機械翻訳モデルの精度の上昇する余地はまだ残っている．

例え連携に適さない参加者と組織化してモデルを統合する場合，学習に参加したデータの総量が増えたとしても，構築されたニューラル機械翻訳モデルの精度が向上するとは限らない．集約するたびに適切な相手と連携すればより高精度のニューラル機械翻訳モデルが得られると推測する．

この推測を検証するために予備実験を行った．まず，従来手法の FedAvg に対して，三つの連邦学習のクライアントを用意する．「Wikipedia 日英京都関連文章対訳コーパス」から「伝統文化」,「歴史」,「人名」の三つのカテゴリから各 70,000 件の対訳を学習データとして，各 10,000 件の対訳を評価データとして，それぞれのクライアントに渡す．各クライアントは OpenNMT でトータル学習ステップが 30,000 のニューラル機械翻訳モデルを構築し．5,000 ステップごとに連邦学習のサーバで計 5 回の集約を行う．各クライアントで構築されたニューラル機械翻訳モデルを使って評価データを翻訳し，BLUE スコアで翻訳精度を測定する．

そして，対照的に節 4.2 で説明するクライアント最適のグリーディ法による自己組織化を対照手法として，従来手法と同じ制約のもとに，集約時に三つのクライアントのモデルを一括に統合するではなく，一部のクライアントのモデルだけが統合されるようにする．

各クライアントが従来手法と対照手法で得たニューラル機械翻訳モデルの翻訳精度は表 1 で示したように，対照手法の方はクライアント A が 35.68 %，ク

表 1: 予備実験で各クライアントが得たニューラル機械翻訳モデルの翻訳精度

	従来手法	対照手法
Client A（伝統文化）	0.14982	0.20327
Client B（歴史）	0.13615	0.15412
Client C（人名）	0.14213	0.18069

ライアント B が 13.20 %，クライアント C が 27.13 %の精度が上昇した．

予備実験で，従来の連邦学習のような全参加者の学習結果を一括に統合する手法より，場合によっては集約時に一部の参加者の学習結果だけが統括される方が構築されたニューラル機械翻訳モデルの翻訳精度が高くなることが証明された．

安定的に精度を上げるために，集約時に特徴量と質が違う学習データを持つ参加者の存在を排除しながら，最適な相手と連携する自己組織化手法が必要である．

第4章　グリーディ法による自己組織化

本章では，グリーディ法を用いて，連邦学習で集約時に連携に適さない参加者の存在を排除しながら，最適な相手と連携する自己組織化の手法について説明する．

グリーディ法（Greedy Algorithm）とは大きな問題をいくつの部分的な問題に分割し，分割された部分的な問題に対する局所最適解を求めることを繰り返す手法である．本研究で提案されたグリーディ法による自己組織化手法は複数の組織が連邦学習でニューラル機械翻訳モデルを生成するという大きな問題をクライアントたちがモデルを訓練してから，サーバでグローバルモデルに集約するまでの流れを部分的な問題に分割して，その集約ごとに生成された新たなグローバルモデルの翻訳精度を最大化にする手法である．これから図 3 とアル

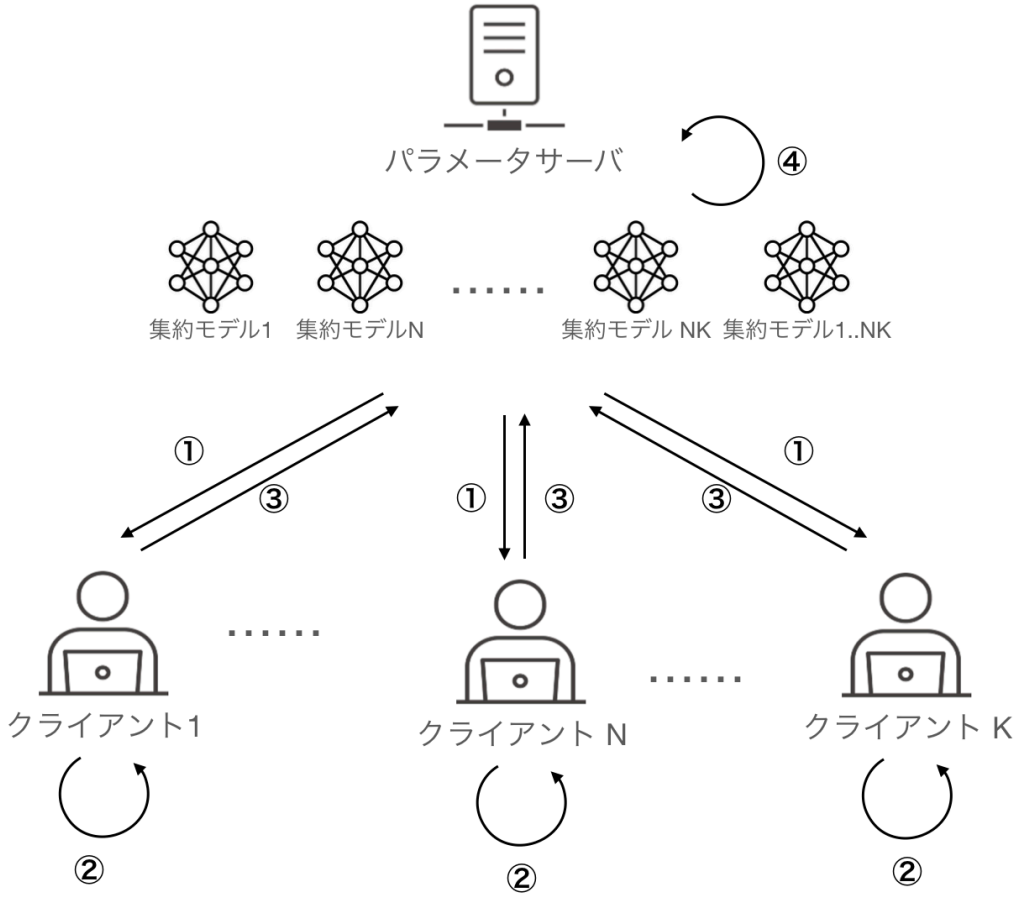

図 3: グリーディ法による自己組織化の流れ

ゴリズム 2 に合わせて具体的な流れ紹介をする．

Algorithm 2 FederatedGreedy.

K はクライアント数; k はクライアントの番号; B はミニバッチサイズ; E は epochs 数; η は学習率; $\nabla l(\omega; b)$ はバッチ b についての損失 $l(\omega; b)$ の勾配

Server executes:

1: Initialize ω_0
2: **for** each round $t = 1, 2, \cdots$ **do**
3: **for** each client k **in parallel do**
4: $\omega_{t+1}^k \leftarrow \text{ClientUpdate}(k, \omega_t)$
5: **end for**
6: Aggregation$(\omega_{t+1}^1, \omega_{t+2}^2, \cdots, \omega_{t+1}^k)$ ▷ モデルを集約
7: **end for**

ClientUpdate(k, ω)**:** ▷ client k で実行

1: $\beta \leftarrow$ (split P_k into batches of size B)
2: **for** each local epoch i from 1 to E **do**
3: **for** batch $b \in \beta$ **do**
4: $\omega \leftarrow \omega - \eta \nabla l(\omega; b)$
5: **end for**
6: return ω to server
7: **end for**

Aggregation$(\omega_{t+1}^1, \omega_{t+2}^2, \cdots, \omega_{t+1}^k)$ ▷ 全ての組み合わせの集約モデルを生成

1: $C \leftarrow$ all combinations of list$[\omega_{t+1}^1, \omega_{t+2}^2, \cdots, \omega_{t+1}^k]$
2: **for** each combination $c \in C$ **do**
3: $\omega_{t+1}^c \leftarrow \sum_{k \in c}^{c} \frac{n_k}{n} \omega_{t+1}^k$
4: **end for**

SelectGlobalModel$(\omega_{t+1}^1, \omega_{t+2}^2, \cdots, \omega_{t+1}^c)$ ▷ グローバルモデルを選出

1: *pass*

---

1. 最初は従来手法と同じようにパラメータサーバは値がランダムに設定されている初期モデルを生成し，各クライアントへ配布する．

2. 各クライアントは，受信したモデルに対して，自身が保有する学習データを用いてローカルで学習を行い，モデルを更新する．
3. 各クライアントは，更新されたモデルからモデルパラメータを抽出し，更新用データとしてサーバへ送信する．
4. クライアントが組み合わせるパターンごとに組織し，サーバはクライアントからもらった更新用データを集約し，パターンごとに集約モデルを作成する．[1)]
5. 作成された集約モデルの中に翻訳精度が最も高いものをグローバルモデルとして選んで，1 のように各クライアントへ配布し，2 の処理に戻る．

複数の組織でニューラル機械翻訳モデルの作成には，どの組織のドメインにも適用できる汎用性が高いニューラル機械翻訳モデルを作成したい時がある．一方，場合によって各組織が自分のドメインに特化したニューラル機械翻訳モデルを期待する時もある．作りたいモデルによって，ステップ５での翻訳精度が最も高いモデルの決め方が変わる．そのため，本研究にはグリーディ法による自己組織化手法をそれぞれの時に特化し，サーバ最適とクライアント最適の二つのパターンを提案した．

4.1 サーバ最適

この節は連邦学習に参加した組織たちがどの組織が管理している対訳コーパスのドメインにも適用できる汎用のニューラル機械翻訳モデルを作成するという目的に特化した手法を説明する．

図 4 のように，クライアント（連邦学習の参加組織）たちはニューラル機械翻訳モデルの翻訳精度を測定するためにデータセットを用意し，サーバと共有する．サーバ側は各クライアントが保有する測定データのサイズの比に合わせて，もらった測定データから共通の測定データを作成する．その後に以下の手順を繰り返す．

まず，サーバは初期モデルを生成し，各クライアントに配布する．

そして，クライアントたちは受信した初期モデルに対して学習し，既定の学習ステップに達したら，学習済みのモデルから集約に必要なパラメータを抽出し，更新用データとしてサーバに返す．

[1)] 例えば三つのクライアントの場合は集約モデル 123，集約モデル 12，集約モデル 13，集約モデル 23，集約モデル 1，集約モデル 2，集約モデル 3，が作成される．

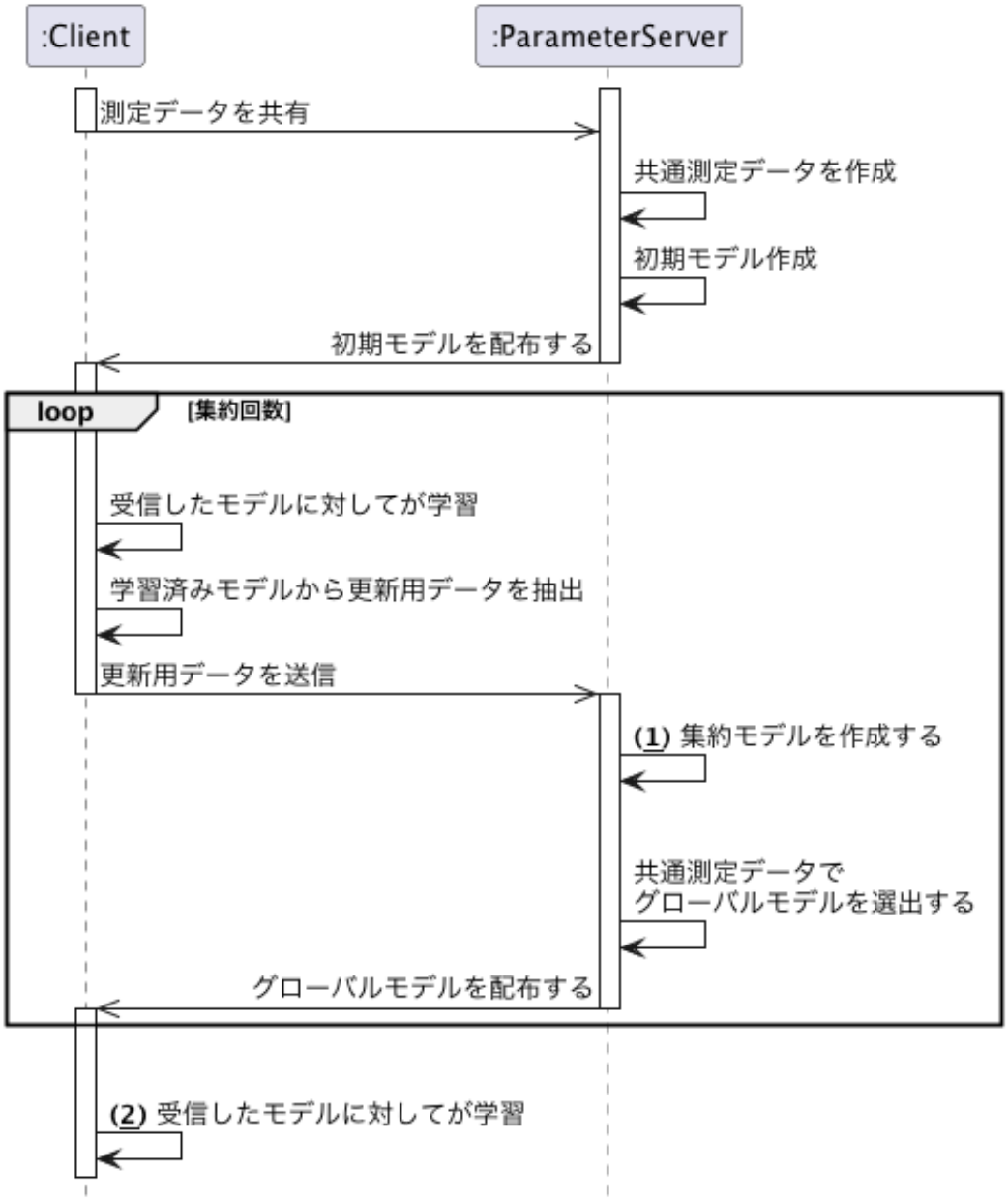

図 4: サーバ最適場合のシーケンス図

次に，サーバ側は更新用データを使って，クライアントが組み合わせるパターンごとに集約モデルを作成する [1)]．

その後，アルゴリズム 3 の SelectGlobalModel に示すようにサーバ側は作成された全ての集約モデルに対して，共通測定データを使って翻訳精度を測定して，精度が最も高い集約モデルをグローバルモデルとして選択され，クライア

[1)] 図 4 の (1)

Algorithm 3 Server optimized FederatedGreedy.

Server executes:

for each round $t = 1, 2, \cdots$ **do**

.....

Aggregation($\omega_{t+1}^1, \omega_{t+2}^2, \cdots, \omega_{t+1}^k$)

SelectGlobalModel($\omega_{t+1}^1, \omega_{t+2}^2, \cdots, \omega_{t+1}^k$) ▷ グローバルモデルを選択

end for

SelectGlobalModel($\omega_{t+1}^1, \omega_{t+2}^2, \cdots, \omega_{t+1}^c$) ▷ サーバで実行

for all $\omega_{t+1}^1, \omega_{t+2}^2, \cdots, \omega_{t+1}^c$ **do**

$\varepsilon \leftarrow$ calculate accuracy from common measurement data

end for

return $\omega_{t+1} \leftarrow \max(\varepsilon)$

ントに配布する．クライアントはモデルに対して保有している学習データで学習し，仕上げる [1)]．

このように，サーバ最適のグリーディ法による自己組織化手法によって，全ての連邦学習に参加した組織が管理している対訳コーパスのドメインにも適応できる汎用性が高いニューラル機械翻訳モデルを得られる．

4.2 クライアント最適

この節は図 5 に合わせて，連邦学習の参加した組織たちが自分の組織が管理している対訳コーパスのドメインに特化したニューラル機械翻訳モデルを作成するという目的に特化した手法を説明する．

サーバ最適と違って，クライアントたちはニューラル機械翻訳モデルの翻訳精度を測定するためのデータセットを用意した後に，サーバに共有する必要がなく，ローカルで次の学習に使うモデルを選択する．

まず，サーバ側に初期モデルを生成し，各クライアントに配布する．

そして，クライアントたちは受信したモデルに対して学習し，既定の学習ステップに達したら，学習済みのモデルから必要なパラメータを抽出し，更新用データとしてサーバに送信する．

[1)] 図 4 の (2)

次に，サーバ側は更新用データをクライアントが組み合わせるパターンごとに集約し，複数の集約モデルを作成[1)]できたら，クライアントにブロードキャストする．

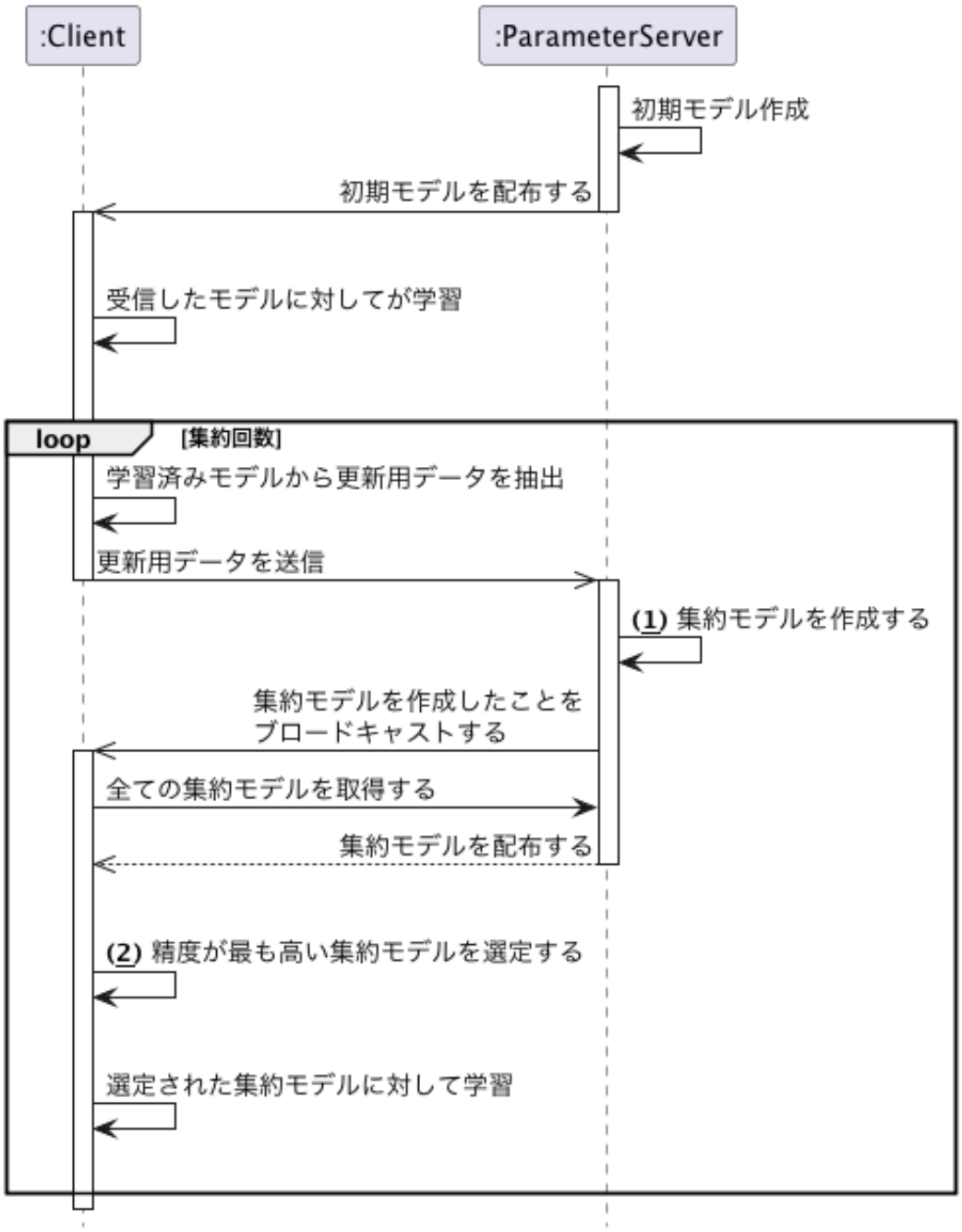

図 5: クライアント最適場合のシーケンス図

[1)] 図 5 の (1)

その後，アルゴリズム 4 が示すようにクライアントたちがサーバから全ての集約モデルを取得し，測定データを使ってローカルで集約モデルの翻訳精度を測定する [1)]．精度が最も高い集約モデルをグローバルモデルとして，次の学習に進む．

Algorithm 4 client optimized FederatedGreedy.

ClientUpdate$(k, \omega_{t+1}^1, \omega_{t+2}^2, \cdots, \omega_{t+1}^c)$**:** ▷ client k で実行

  $\omega \leftarrow$ SelectGlobalModel$(\omega_{t+1}^1, \omega_{t+2}^2, \cdots, \omega_{t+1}^c)$ ▷ グローバルモデルを選出

  $\beta \leftarrow$ (split P_k into batches of size B)

  for each local epoch i from 1 to E **do**

    ……

  end for

SelectGlobalModel$(\omega_{t+1}^1, \omega_{t+2}^2, \cdots, \omega_{t+1}^c)$ ▷ client k で実行

  for all $\omega_{t+1}^1, \omega_{t+2}^2, \cdots, \omega_{t+1}^c$ **do**

    $\varepsilon \leftarrow$ calculate accuracy from measurement data

  end for

  return $\omega \leftarrow \max(\varepsilon)$

サーバで最後の集約を行った後に，クライアントたちはサーバから全ての集約モデルを取得し，精度が最も高い集約モデルに対して，保有している学習データを使って学習し，自ドメインでの翻訳精度を上昇するように仕上げる．

サーバ最適と比べ，クライアント最適の場合は測定データをサーバに共有しないため，より高いセキュリティ性が得られる上に，集約モデルを選定する時に測定データのサイズ，特徴量などを調整することで，より自ドメインに適するニューラル機械翻訳モデルを期待できる．

このように，クライアント最適のグリーディ法による自己組織化手法によって，それぞれの連邦学習に参加した組織に属する分野に特化したニューラル機械翻訳モデルを構築できる．

[1)] 図 5 の (1)

第5章　マルチエージェント深層強化学習による自己組織化

連邦学習に複数の参加者がいるため，参加者たちのモデルを組み合わせるパターンが多様である．ある組み合わせで得たニューラル機械翻訳モデルの精度が一時的に下がっても，最終的の精度が上昇することもありうる．また，連邦学習の参加者の増加と共に組み合わせるパターンも増大する．その結果，膨大な組み合わせパターンで最善な組織法を見つけ出すことが困難となる．

本章では，マルチエージェント深層強化学習を用いて，将来的なモデル精度の期待値を考慮し，組織化する時のコストを抑える自己組織の手法について説明する．

5.1　マルチエージェント深層強化学習

強化学習（Reinforcement Learning）は機械学習の一種である．心理学の中に刺激と応答の間に与えられる刺激（強化子）を元に特定の行動パターンの発見が増強されることを強化と呼ぶ．図 6 のように強化学習にはある環境の中にエージェントがいると考える．エージェントが現在の状態を観測し，次に取り込む行動を決定する．その行動を行うことで環境から報酬を得られる．報酬の大きさを刺激の強さにみなし，一連の行動を通じて刺激がもっとも強くになるようなポリシーを学習する．

Q-Learning は代表的な強化学習手法である．[6] 強化学習の中にある状態 s において，とある行動 a を取った時の価値を Q 値で表す．ある状態 s にとあ

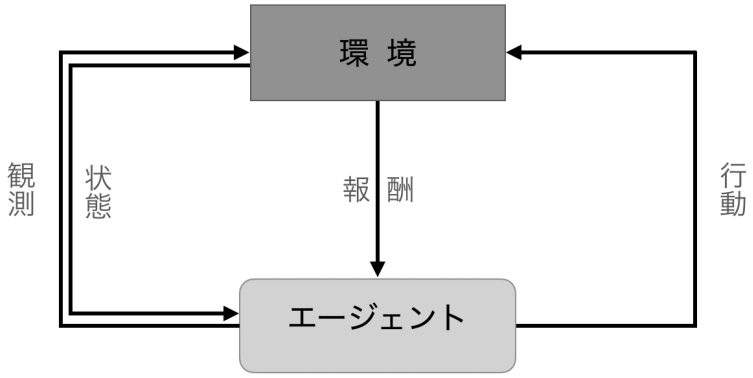

図 6: 典型的な強化学習の構成

る行動 a を取った時の価値は Q であるという情報を Q テーブルで保持する．Q-Learning では，式（1）のようにエージェントが一つの行動を取った段階（1 STEP）で得た報酬 r のもとに現状態のもとに取った行動の価値 Q を見積もり，Q テーブルを更新する．

$$Q^{new}(s_t, a_t) \leftarrow (1-a)Q(s_t, a_t) + \alpha[r_{t+1} + \gamma \max a_{t+1} Q(s_{t+1}, a_{t+1})]^{1)} \quad (1)$$

Q テーブルには大きな問題を抱え込んでいる．それは状態行動空間の爆発と呼ばれるものである．状態や行動が連続値，またはその次元が非常に大きい環境において，Q テーブルを保持するために，無限の領域が必要になり，現実的な考えではない．

この問題を解決したのは深層強化学習である．深層強化学習は状態空間の爆発に対してニューラルネットワークを用いて Q テーブルを関数近似する手法である．また，Q-Learning ベースの深層強化学習は DQN（Deep Q-Network）と呼ばれる．[7][8]

しかしながら，実世界にとある環境の中に複数のエージェントが存在るのは

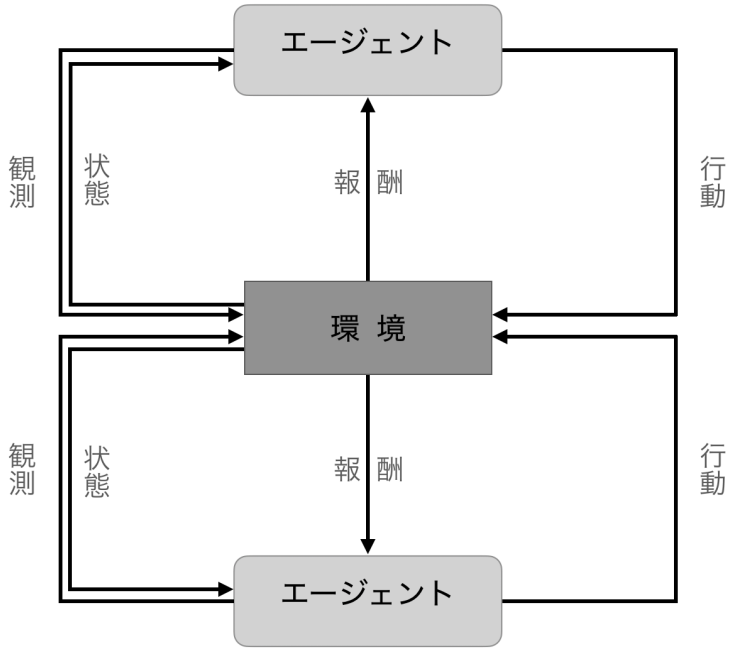

図 7: マルチエージェント強化学習の構成

[1)] α は学習率（$0 \leq \alpha \leq 1$），γ は割引率（$0 \leq \gamma \leq 1$），t はとある時点

普遍である．図 7 のように，同じ環境で複数の強化学習エージェントが同時に学習，行動し，相互影響する自律分散型システムはマルチエージェント強化学習という．

マルチエージェント強化学習にはエージェントの利害関係によって以下の種類がある．一つ目は全てのエージェントが協力し，システム全体の報酬を最大化にする完全協力型（Fully Cooperative）．二つ目はあるエージェントが勝利すると他のエージェントが負けになる完全競争型（Fully Competitive）．三つ目はエージェントたちに競争と協力の関係が同時に存在する混合：協力＆競争型（Mixed Cooperative Competitive）．四つ目はエージェントが自分の利益だけを最大化にする利己型（Self-interested）．

また，学習方法によって三つに分類できる．シングルエージェントの場合と同様中央集権的なエージェントが他のエージェントの学習，行動をコントロールする完全中央集権型（Fully Centralized）．各エージェントが独立して学習，行動を決定する完全非中央集権型（Fully Decentralized）．中央集権的なエージェントに学習して，他エージェントが行動を決定する混合型（Mixed：Centralized Decentralized）．

この三つの学習方法の中，図 8 のように，完全中央集権型の学習が最も安定し，収束しやすくに対して，エージェント間の相互作用を考えないといけないため学習空間が大きくなる傾向がある．一方，完全非中央集権型にはエージェント間の相互作用を考えなくでもいいため，学習空間が小さくで済むだが，その分エージェントたちが互いの状態を把握しきれず，学習が収束しにくく，不

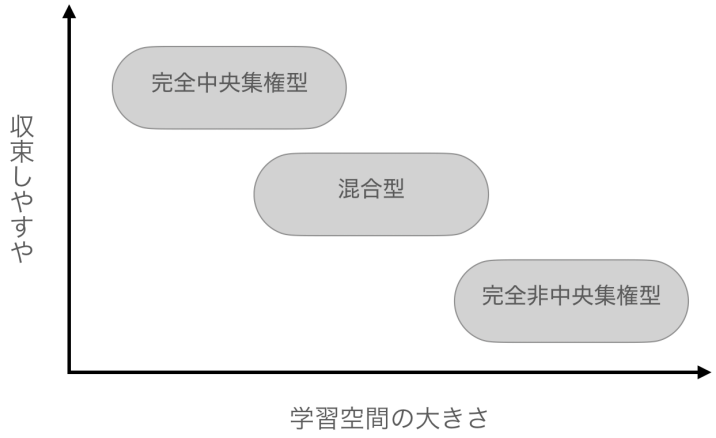

図 8: マルチエージェント強化学習学習方法の比較

安定になる傾向がある．

5.2 モデル化

マルチエージェント深層強化学習を用いて，将来的なモデル精度の期待値を考慮し，最善な組織法を見つけ出せる．これを実現するために，連邦学習システムをモデル化し，マルチエージェント深層強化学習システムに取り込む必要がある．図 9 に合わせてマルチエージェント深層強化学習を用いた連邦学習システムを説明する．

環境

連邦学習システム（FedLearning）をマルチエージェント深層強化学習の環境と設定する．強化学習エージェントが連邦学習のクライアントを通じて環境である連邦学習システムを観測したり，影響したりする．環境の中に複数の連邦学習クライアント＆強化学習エージェントのペアと一つの連邦学習サーバが存在する．

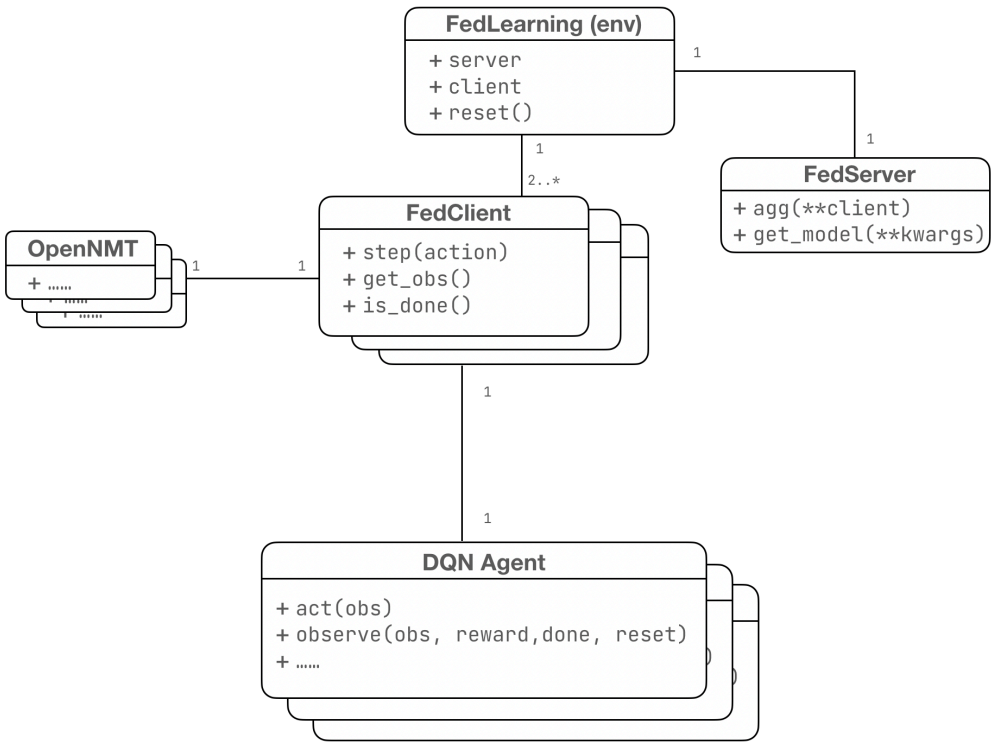

図 9: マルチエージェント深層強化学習を用いた連邦学習の構成

サーバ

サーバ（FedServer）は連邦学習システムのサーバのことを指す．環境を初期化する時にサーバは初期モデルを生成し，クライアントに配る役割がある．また，クライアントからリクエストが来たら，リクエスト通りに指定されたクライアントたちを一つの組織になり，モデルを集約し，集約したニューラル機械学習モデルをリクエストしたクライアントに返す役割も持っている．

クライアント

クライアント（FedClient）は連邦学習システムのクライアントのことを指す．環境の初期化する時にクライアントはサーバから受信した初期モデルに対して学習し，モデルが集約できるようにする．クライアントはサーバに誰と組織になるかのをリクエストでき，サーバからもらった集約されたモデルに対してさらなる学習することができる．また，クライアントは環境の状態を観測したり，報酬を算出する役割も持っている．

エージェント

エージェント（DQN Agent）はマルチエージェント深層強化学習システムのエージェントのことを指す．一個のクライアントに一つのエージェントを付き，ペアで存在する．エージェントはペアになっているクライアントを通じて，環境を観測し，現在の状態 s を取得する．エージェントが現在の状態 s と自分が持っているニューラル強化学習モデルで次の行動 a を選択する．決めた行動 a に基づいてクライアントがどのようなリクエストを送るかをコントロールする．また，ペアになっているクライアントが報酬 r を算出したら，現在の状態 s，取った行動 a，得た報酬 r で式 1 のように Q 値を算出し，ニューラル強化学習モデルを更新する．

また，図 10 が示したように，エージェントはペアになっているクライアントだけに対して学習し，他のエージェントの状態や行動などを無視する完全非中央集権型である．

状態

強化学習エージェントがクライアントに通じて環境を観測し，現在の状態 s を取得する．クライアントは学習済みニューラル機械翻訳モデルを使って，事前に用意した測定データの各センテンスを翻訳し，翻訳精度を測定する．式（2）のように各センテンスの翻訳精度を保持する行列が現在の状態を表す．また，環境を初期化し，クライアントが初期モデルに対して学習し終わった時の状態を

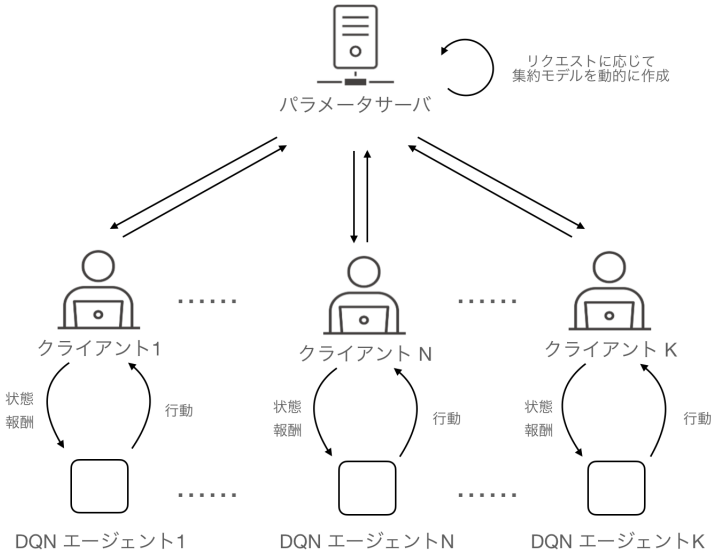

図 10: マルチエージェント深層強化学習を用いた連邦学習システムの概要

初期状態という．

$$s^{now} = [score(sentence_0), score(sentence_1), \cdots, score(sentence_k)]^{1)} \quad (2)$$

行動

強化学習エージェントは状態 s のもとに自分が保持しているニューラル強化学習モデルを通じで次の行動 a を決定する．行動 a はどのクライアントとどのクライアントが組織になって，集約されたニューラル機械翻訳モデルに対して既定の学習ステップで学習するかのことである．組織の中に複数のクライアントがいても構わないが，単体のクライアントで組織になることもあり得る．組織になるクライアントは自分とペアしているクライアントを含まなくでも良いと考える．すなわち，クライアント数が m の時，行動の種類は $\sum_{n=1}^{m} \frac{m!}{n!(m-n)!}$ となる．

報酬

強化学習エージェントが行動を取り，環境を影響した結果，環境から報酬を受け取る．報酬 r はクライアントの前回の学習済みニューラル機械翻訳モデルから今回の学習済みニューラル機械翻訳モデルの翻訳精度の上昇量である（$r = score_t - score_{t-1}$，$-1 \leq r \leq 1$）．ニューラル機械翻訳モデルの精度（$score_t$）は測定データの各センテンスの翻訳精度の平均で算出する．式で表

[1)] k は測定データのセンテンス数

すと $score = \bar{s} = \overline{score(sentence_0), score(sentence_1), \cdots, score(sentence_k)}$ である．

マルチエージェント深層強化学習を用いた連邦学習システムに複数のニューラルモデルが存在し，三種類のが学習が行われて，複雑である．そのため，図 11 に合わせて，マルチエージェント深層強化学習による自己組織化手法の流れを説明する．

まず，連邦学習単体と同様にサーバは値がランダムに設定されている初期モデルを生成し，各クライアントに配布する．各クライアントは受信した初期モデルに対して，自身が保有する学習データを用いてローカルで既定のステップの学習を行う．

そして，クライアントたちは学習を済んだら，学習済みニューラル機械翻訳モデルからモデルパラメータを抽出し，更新用データとしてサーバへ送信する．同時に，クライアントたちは各自が持つ測定データを使って，学習済みニューラル機械翻訳モデルの翻訳精度 $score_0$ を測定し，初期状態 s_0 を得る．

次に，それぞれのクライアントはペアの強化学習エージェントに初期状態 s_0 を渡す．エージェントは状態 s_0 と自分が所持するニューラル強化学習モデルを通じて行動 a_0 を選択し，クライアントに伝える．

クライアントは行動 a_0 にしたがって，サーバにリクエストする．サーバ側はリクエストに応じてクライアントたちを組織し，更新用データを集約して，集約されたニューラル機械翻訳モデルをレスポンスとしてクライアントに返す．

各クライアントは自分がリクエスした集約モデルを受信した後に，それに対してローカルで既定のステップになるまで学習し続ける．学習し終わったら，学習済みのニューラル機械翻訳モデルから更新用データを抽出しサーバに送信するの同時に，測定データで学習済みのニューラル機械翻訳モデルの翻訳精度 $score_1$ と現在の状態 s_1 を測り，エージェントに与える報酬（$r_0 = score_1 - score_0$）を算出する．

強化学習エージェントは現在の状態 s_1 と報酬 r_0 を使って式（1）で $Q(s_0, a_0)$ を更新し，自分が持つニューラル強化学習モデルに学習させ，次の行動 a_1 を選択する．そのあと，図 11 の Loop の部分のように繰り返す．

最後はクライアントたちはエージェントの行動で得たニューラル機械翻訳の集約モデルを自分のグローバルモデルとして保存し，学習を終了する．

このように，マルチエージェント深層強化学習による自己組織化手法は試行錯誤の中で強化学習エージェントが最善な連邦学習のクライアントの組織法を身につけ，構築されたニューラル機械翻訳モデルの精度を最大限に引き上げる．

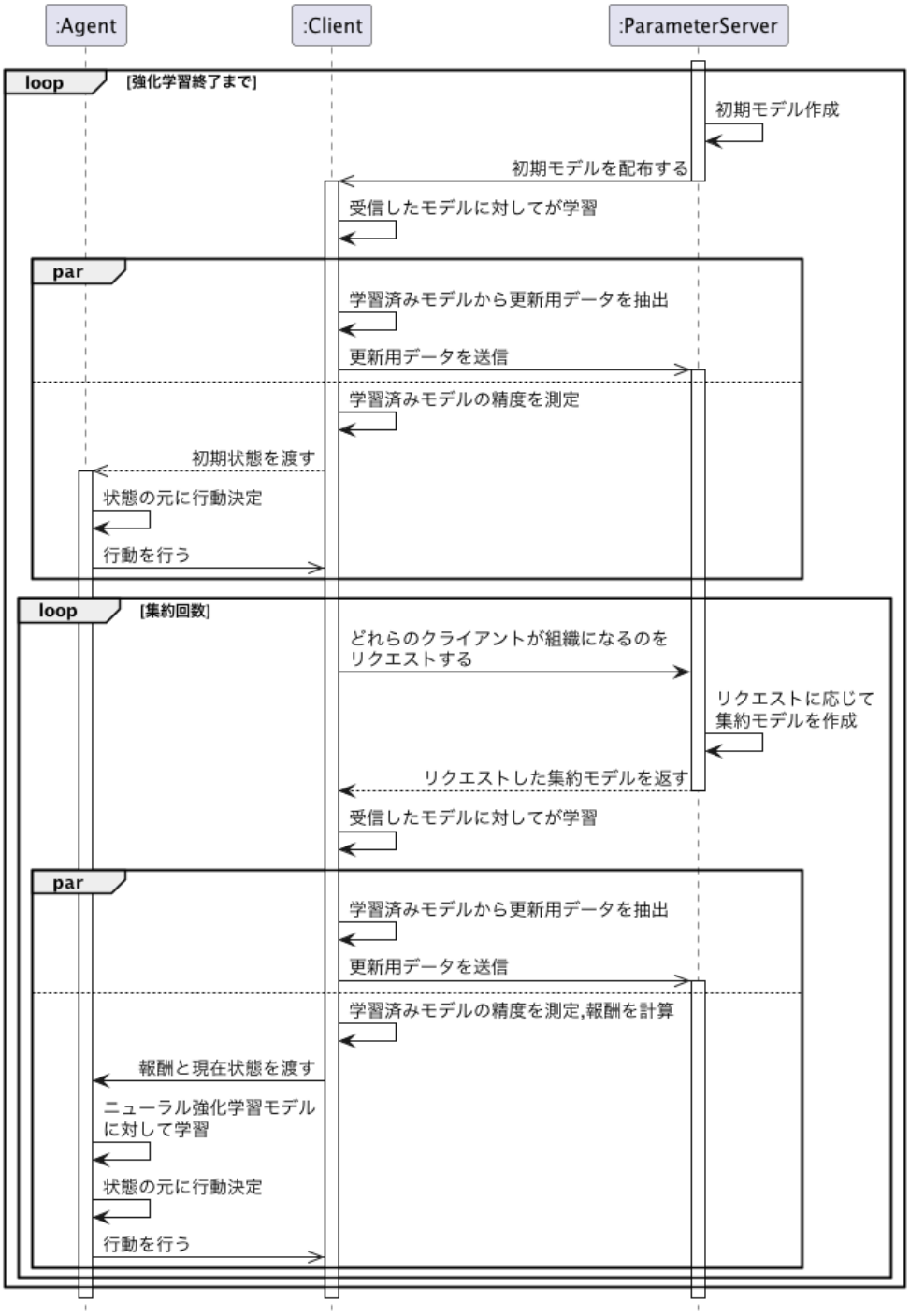

図 11: マルチエージェント深層強化学習による自己組織化手法のシーケンス図

第 6 章　評価

本章では，はじめに実験環境を説明する．次に，それぞれの手法を比較する際に使用する評価指標について説明する．続いて，グリーディ法による自己組織化手法とマルチエージェント深層強化学習による自己組織化を実験し，提案手法の有効性の確認を行う．最後に実験結果について考察する．

6.1　評価方法

6.1.1　実験環境

実際に評価を行った検証システムについて説明する．提案手法の有用性を検証するために，三つの組織が参加している連邦学習システムを考案し，構築した．それぞれの組織が一個の連邦学習クライアントを持ち，組織たちが共同で連邦学習サーバを維持する．図 12 に構成図を示す．ニューラル機械翻訳モデルの構築は OpenNMT[9] というニューラル機械翻訳システムを利用する．

そして，対訳コーパスは「Wikipedia 日英京都関連文書対訳コーパス」を使用する．「Wikipedia 日英京都関連文書対訳コーパス」は国立研究開発法人情報通信研究機関が Wikipedia の京都関連の日本語記事に対して人手で翻訳し，作成した精密かつ大規模なコーパスである．Wikipedia 日英京都関連文書対訳コー

図 12: 検証システム構成

パス」は「建造物」,「神道」,「仏教」など計１５のカテゴリに分割される．本研究は「伝統文化」,「歴史」,「人名」の三つのカテゴリをそれぞれを 80,000 件の対訳データを抽出し，データセットとして各クライアントに管理させる．表 2 各データセット間のコサイン類似度を示す. 各クライアントは自分のデータセットから 70,000 件をニューラル機械翻訳モデルを構築するための学習データとして分け，残りの対訳データをクライアントたちが動的組織化する時にモデルの精度を測るための測定データに割り当てる（表 3）．ニューラル機械翻訳モデルの精度は BLEU スコアで算出する．

各クライアントはニューラル機械翻訳モデルに対して 5,000 ステップを学習するごとに，サーバで 1 回の連邦学習の集約を行う．各クライアントはニューラル機械翻訳モデルに対してバッチサイズを 32 ，トータル学習回数は 30,000 ステップに設定し，5 回の集約が行われる（5,000 ステップごとに集約が行われる）.

6.1.2 評価指標

各手法の有効性を評価する指標として，従来手法である FedAvg によりニューラル機械翻訳モデルを構築し，翻訳精度を測定する．

評価データと共通評価データを通じて，構築されたニューラル機械翻訳モデルの翻訳精度の評価を行う．評価データは測定データと同じものであり，構築

表 2: 各データセット間のコサイン類似度

	伝統文化	歴史	人名
伝統文化	1	0.59339	0.82510
歴史	-	1	0.81514
人名	-	-	1

表 3: 各クライアントのデータセット

	カテゴリ	学習データ	測定データ
Client A	伝統文化	70,000	10,000
Client B	歴史	70,000	10,000
Client C	人名	70,000	10,000

表 4: FedAvg により構築したニューラル機械翻訳モデルの翻訳精度

	カテゴリ	精度 (自分野)	精度 (全分野)
Client A	伝統文化	0.14982	0.13859
Client B	歴史	0.13615	0.16081
Client C	人名	0.14213	0.14806

されたニューラル機械翻訳モデルが自分野 [1)]での精度（専門性）を評価する時に使用する．

FedAvg によるニューラル機械翻訳モデルの構築は節 6.1.1 で説明した検証システムを使って，同じ環境設定のもとで行う．表 4 は得たニューラル機械翻訳モデルのの学習データのカテゴリと自分野，全分野の精度（BLEU スコア）である．これより，提案手法により構築されたニューラル機械翻訳モデルの自分野と他分野の精度の上昇量を算出できる．共通評価データは三つのクライアントが持つ測定データをシャッフルし，３分割して作られるものであり，構築されたニューラル機械翻訳モデルが全分野での精度（汎用性）を評価する時に使用する．

また，三つのデータセットを一つのクライアンに集中し，そのクライアントで中央集権型のニューラル機械翻訳モデルを構築する．構築された中央集権型のニューラル機械翻訳モデルの精度は今回の実験で使用しているデータセットで得られるニューラル機械翻訳モデルの最大精度にみなす．精度を比較するために，中央集権型で学習する際の epoch 数を連邦学習で学習する時の epoch 数に一致する必要がある．式 3 は epoch の計算式である．

$$epoch = train\ steps \div \frac{data\ size}{batch\ size} \tag{3}$$

そのため，中央集権型での学習ステップは式 4 によって 90,000 である．構築されたニューラル機械翻訳モデルの精度は表 5 で示す．これより，提案手法により構築されたニューラル機械翻訳モデルの精度は最大精度との割合を算出できる．

[1)] 例えば Client A の自分野は伝統文化である．

$$
\begin{aligned}
epoch_{fed} &= epoch_{centralized} \\
30000 \div \frac{70000}{32} &= train\ steps_{centralized} \div \frac{210000}{32} \\
train\ steps_{centralized} &= 90000
\end{aligned}
\tag{4}
$$

6.2 グリーディ法の評価結果

6.2.1 サーバ最適

節 4.1 で説明した通り，サーバ最適ための共通測定データを作成し，ニューラル機械翻訳モデルを構築する．集約時にクライアントたちはグリーディ法による自己組織法に基づいて組織を形成する．各集約時に形成された組織の履歴は表 6 である．得られたニューラル機械翻訳モデルの測定精度は図 13 のように遷移する．

構築されたニューラル機械翻訳モデルを独自評価データで評価した結果は表 7 に示す．図 14 は独自評価データで評価した結果と従来手法の比較である．自分野において，サーバ最適の時にグリーディ法による自己組織化を用いて，Client A が構築したニューラル機械翻訳モデルが従来手法で構築されたニューラル機械翻訳モデルの精度より 43.28 %上昇した．Client B の場合は 15.69 %上昇した．Client C の場合は 26.67 %上昇した．

サーバ最適の時にグリーディ法による自己組織化で構築されたニューラル機械翻訳モデルの専門性が従来手法より高いと言える．

表 5: 中央集権型のニューラル機械翻訳モデルの精度

カテゴリ	学習データ	測定データ	精度
伝統文化, 歴史, 人名	90,000	30,000	0.20234

表 6: サーバ最適時にグリーディ法によるクライアント自己組織履歴

	1回目	2回目	3回目	4回目	5回目
Client A	A	C	AC	ABC	ABC
Client B					
Client C					

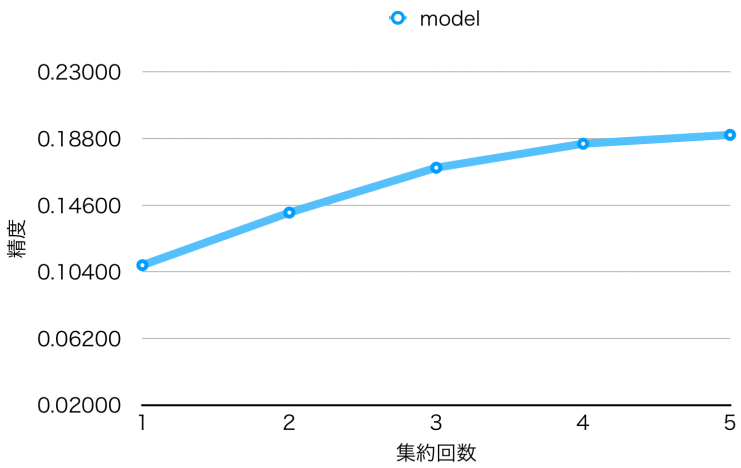

図 13: サーバ最適時に各集約時にニューラル機械翻訳モデルの精度遷移

構築されたニューラル機械翻訳モデルを共通評価データで評価した結果は表 8 に示す．図 15 は共通評価データで評価した結果と従来手法の比較である．全分野において，サーバ最適の時にグリーディ法による自己組織化を用いて， Client A が構築したニューラル機械翻訳モデルが従来手法により構築されたニューラル機械翻訳モデルの精度が 28.41 %上昇した．Client B の場合は 6.66 %上昇した．Client C の場合は 17.03 %上昇した．

サーバ最適の時にグリーディ法による自己組織化で構築されたニューラル機械翻訳モデルの汎用性が従来手法より高いと言える．

表 7: 独自評価データでサーバ最適時にグリーディ法の評価結果

	学習データ	測定データ	評価データ	精度
Client A	伝統文化	共通	伝統文化	0.21467
Client B	歴史		歴史	0.15751
Client C	人名		人名	0.18004

表 8: 共通評価データでサーバ最適時にグリーディ法の評価結果

	学習データ	測定データ	評価データ	精度
Client A	伝統文化	共通	共通	0.17797
Client B	歴史			0.17152
Client C	人名			0.17327

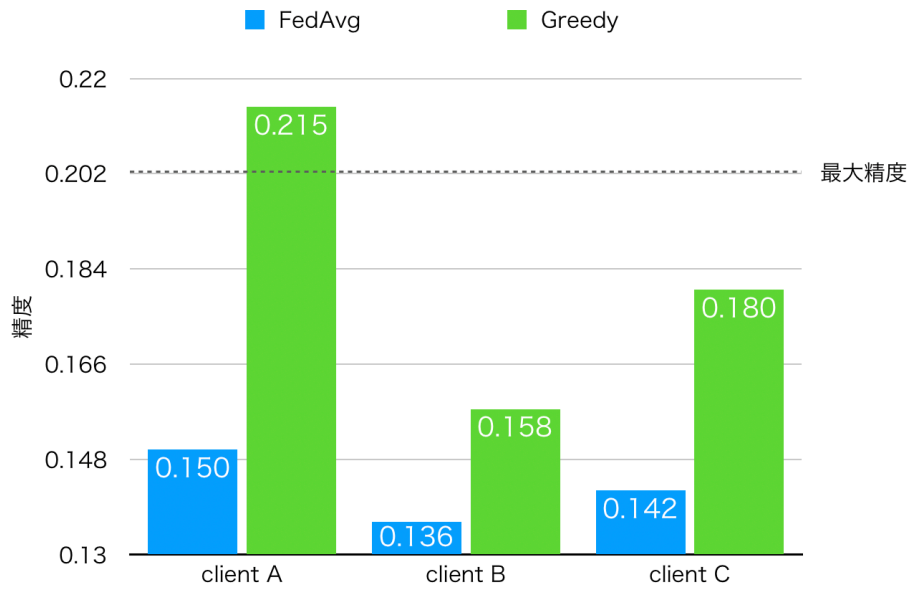

図 14: 独自評価データでサーバ最適時にグリーディ法と従来手法の比較

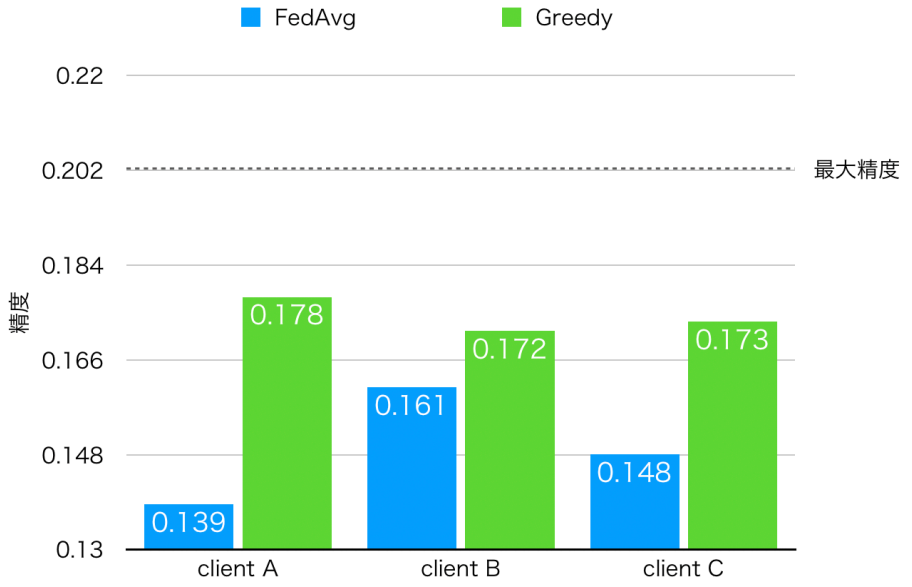

図 15: 共通評価データでサーバ最適時にグリーディ法と従来手法の比較

6.2.2 クライアント最適

節 4.2 で説明した通りにニューラル機械翻訳モデルを構築する．各集約時に形成された組織の履歴は表 9 である．得られたニューラル機械翻訳モデルの測定精度は図 16 のように遷移する．

構築されたニューラル機械翻訳モデルを独自評価データで評価した結果は表

表 9: クライアント最適時にグリーディ法によるクライアント自己組織履歴

	1 回目	2 回目	3 回目	4 回目	5 回目
Client A	A	AC	A	A	A
Client B	A	C	BC	BC	BC
Client C	A	C	BC	BC	BC

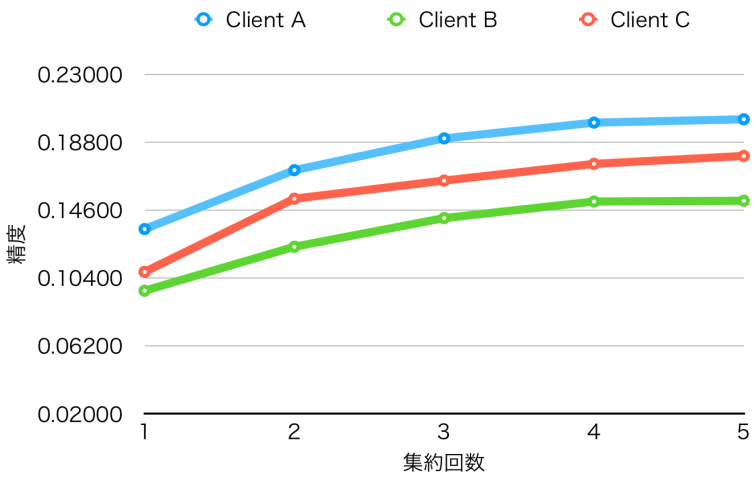

図 16: クライアント最適時に各集約時にニューラル機械翻訳モデルの精度遷移

表 10: 独自評価データでクライアント最適時にグリーディ法の評価結果

	学習データ	測定データ	評価データ	精度
Client A	伝統文化	伝統文化	伝統文化	0.20327
Client B	歴史	歴史	歴史	0.15412
Client C	人名	人名	人名	0.18069

10 に示す．図 17 は独自評価データで評価した結果と従来手法の比較である．自分野において，クライアント最適の時にグリーディ法による自己組織化を用いて，Client A が構築したニューラル機械翻訳モデルが従来手法で構築されたニューラル機械翻訳モデルの精度より 35.68 %上昇した．Client B の場合は 13.20 %上昇した．Client C の場合は 27.13 %上昇した．

クライアント最適の時にグリーディ法による自己組織化で構築されたニューラル機械翻訳モデルの専門性が従来手法より優れている．

構築されたニューラル機械翻訳モデルを共通評価データで評価した結果は表 11

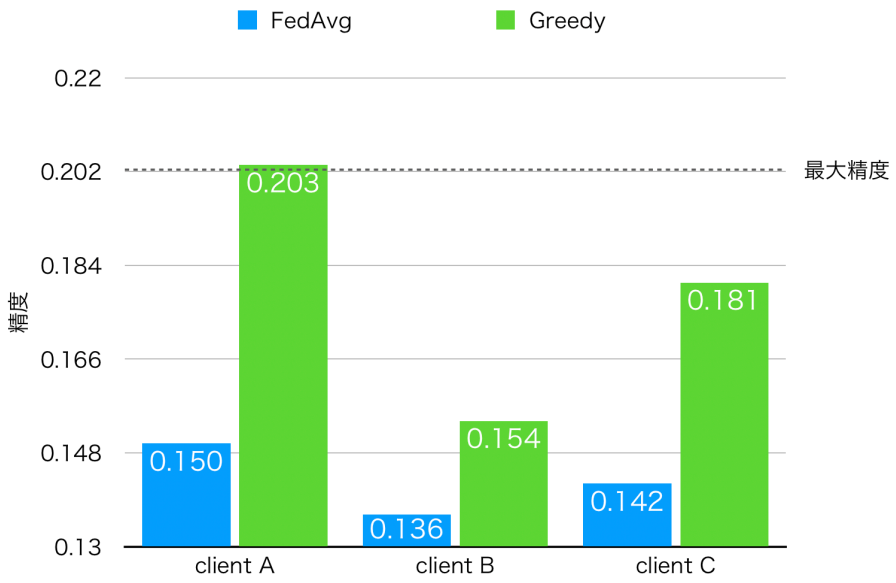

図 17: 独自評価データでクライアント最適時にグリーディ法と従来手法の比較

表 11: 共通評価データでクライアント最適時にグリーディ法の評価結果

	学習データ	測定データ	評価データ	精度
Client A	伝統文化	伝統文化	共通	0.16270
Client B	歴史	歴史		0.16186
Client C	人名	人名		0.17041

に示す．図 18 は共通評価データで評価した結果と従来手法の比較である．全分野において，サーバ最適の時にグリーディ法による自己組織化を用いて， Client A が構築したニューラル機械翻訳モデルが従来手法より構築されたニューラル機械翻訳モデルの精度より 17.40% 上昇し，Client B が構築したモデルは従来手法より 0.65% 高く，Client C の場合は 15.10% 上昇した．

クライアント最適の時にグリーディ法による自己組織化で構築されたニューラル機械翻訳モデルの汎用性が従来手法より優れている．

6.3 マルチエージェント深層強化学習法の評価結果

マルチエージェント強化学習による自己組織化手法を評価するために，節 6.1.1 に説明した検証システムのもとに，図 19 各組織に一個の深層強化学習エージェント（DQN Agent）を追加する．深層強化学習エージェントは PFRL という

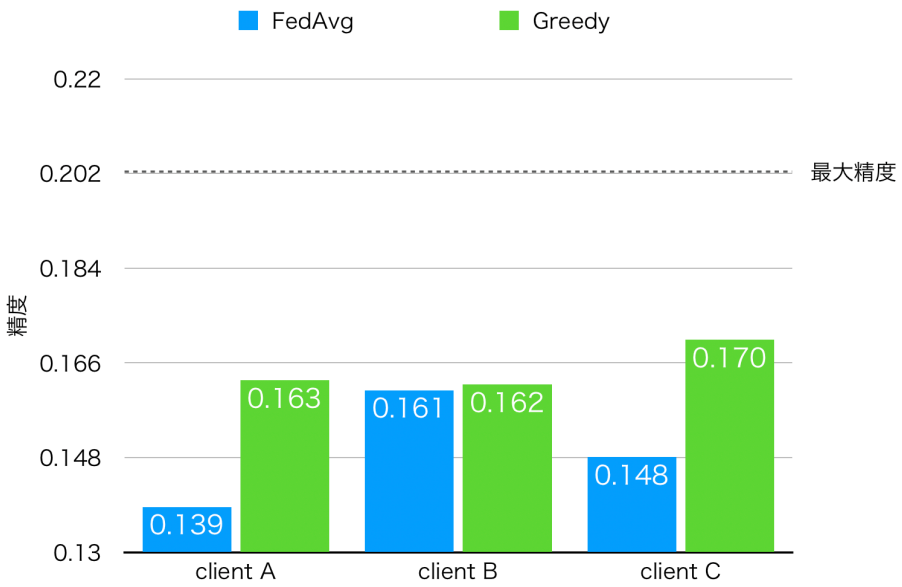

図 18: 共通評価データでクライアント最適時にグリーディ法と従来手法の比較

PyTorch 向けの深層強化学習ライブラリを利用して実装された．

各深層強化学習エージェントは同一組織のクライアントを通じて状態を観察するため，状態空間のサイズは各クライアントが保有している測定データの件数と同じく 10,000 である．また，検証システムに三つの組織が存在しているた

図 19: マルチエージェント強化学習による自己組織化手法ための検証システム構成

め，深層強化学習エージェントの行動空間のサイズは $\sum_{n=1}^{3} \frac{3!}{n!(3-n)!} = 7$ である．Q 値を予測するために三層の強化学習ニューラルネットワークを使用し，入力層のノードは状態空間にあわせて 10,000 個，出力層のノードは行動空間に従って 7 個がある．

各深層強化学習エージェントが新しい Q 値を探索（exploration）する際に $\epsilon - greedy$ アルゴリズムに基づき，経験の割引率 γ を 0.9 にする．

図 20 のように，強化学習の中に連邦学習で 5 回の集約を行なって，各クライアントが 30,000 ステップの学習を完了し，ニューラル機械翻訳モデルを構築できたら 1 エピソードとしてカウントする．

図 11 と図 20 が示すように提案手法に従って，連邦学習でニューラル機械翻訳モデルを構築し，マルチエージェント深層強化学習に 290 エピソードを学習させる．マルチエージェント深層強化学習の各エピソードで構築されたニューラル機械翻訳モデルの精度を記録し，図 21 の折れ線グラフを作成した．また，区間を 20 エピソードで各クライアントが構築されたニューラル機械翻訳モデルの精度の移動平均を計算し，図 21 に濃い色の線で表現した．移動平均線から深層強化学習エージェントが決めたクライアントの組織パターンにより構築されたニューラル機械翻訳モデルの精度が徐々に上がることが一目瞭然である．深層強化学習エージェントが試行錯誤を通じて適切なクライアントを組織になる能力が上達になっている．

また，精度が最も高いニューラル機械翻訳モデルを構築できるクライアンの組織パターンを最善な組織パターンと決める．今回の実験で，マルチエージェント深層強化学習による自己組織法で見つけた最善な組織パターンは表 12 に示す．最善な組織パターンで構築したニューラル機械翻訳の精度は表 13 に記載する．図 22 は独自評価データで最善な組織パターンで構築したニューラル機械翻訳モデルとクライアント最適のグリーディ法による自己組織化手法や従来手法で構築したニューラル機械翻訳モデルの評価結果の比較である．

自分野において，最善な組織パターンで Client A が構築したニューラル機械翻訳モデルの精度は従来手法で構築されたニューラル機械翻訳モデルの精度より 46.98% 上昇した．Client B が構築したニューラル機械翻訳モデルの精度は従来手法で構築されたニューラル機械翻訳モデルの精度より 19.76% 上昇した．Client C が構築したニューラル機械翻訳モデルの精度は従来手法で構築された

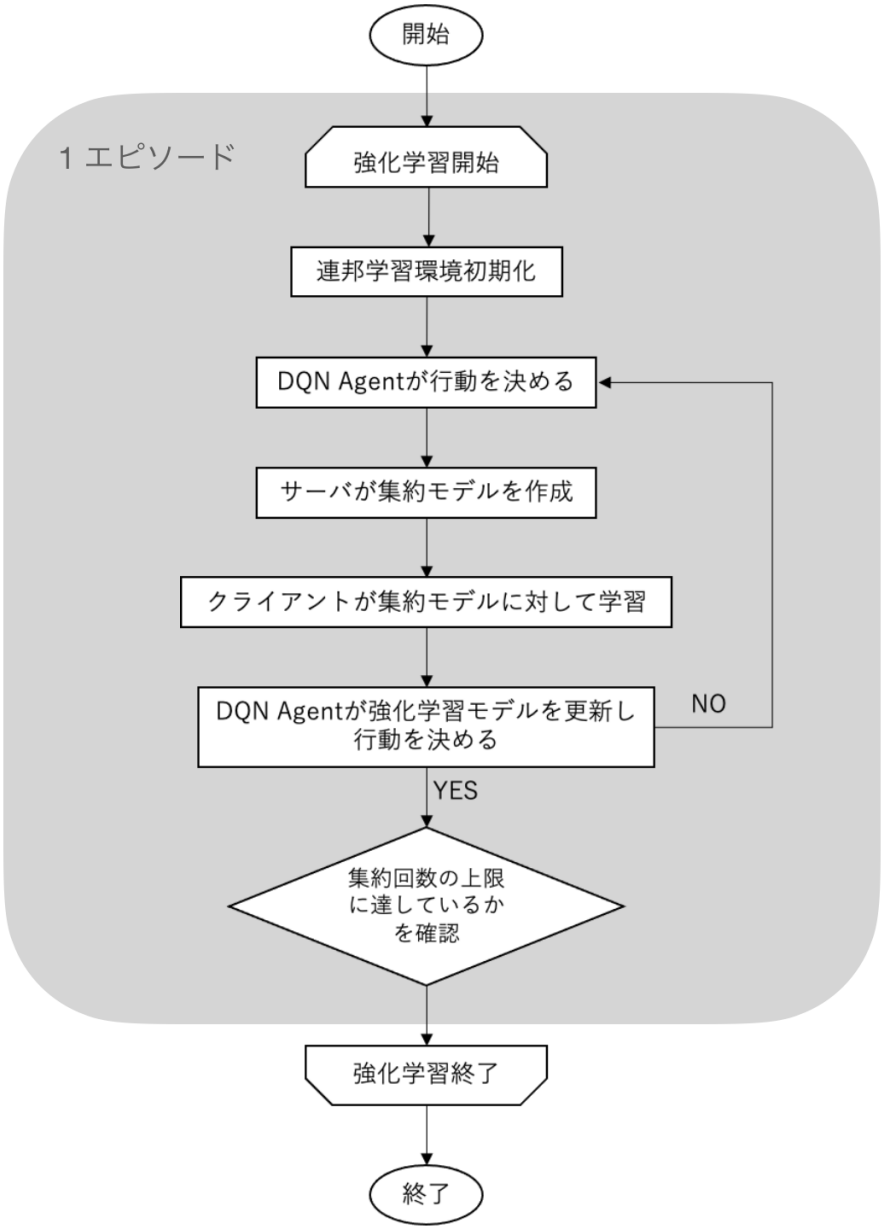

図 20: 検証時に強化学習の流れ

ニューラル機械翻訳モデルの精度より 32.91% 上昇した.

6.4 考察

本節では，節 6.2 と節 6.3 で提示した結果のもとに二つの提案手法で構築したニューラル機械翻訳モデルの精度について考察を与えた後，それぞれの提案手法の問題点を指摘し，改善法を提案する．最初にグリーディ法による自己組

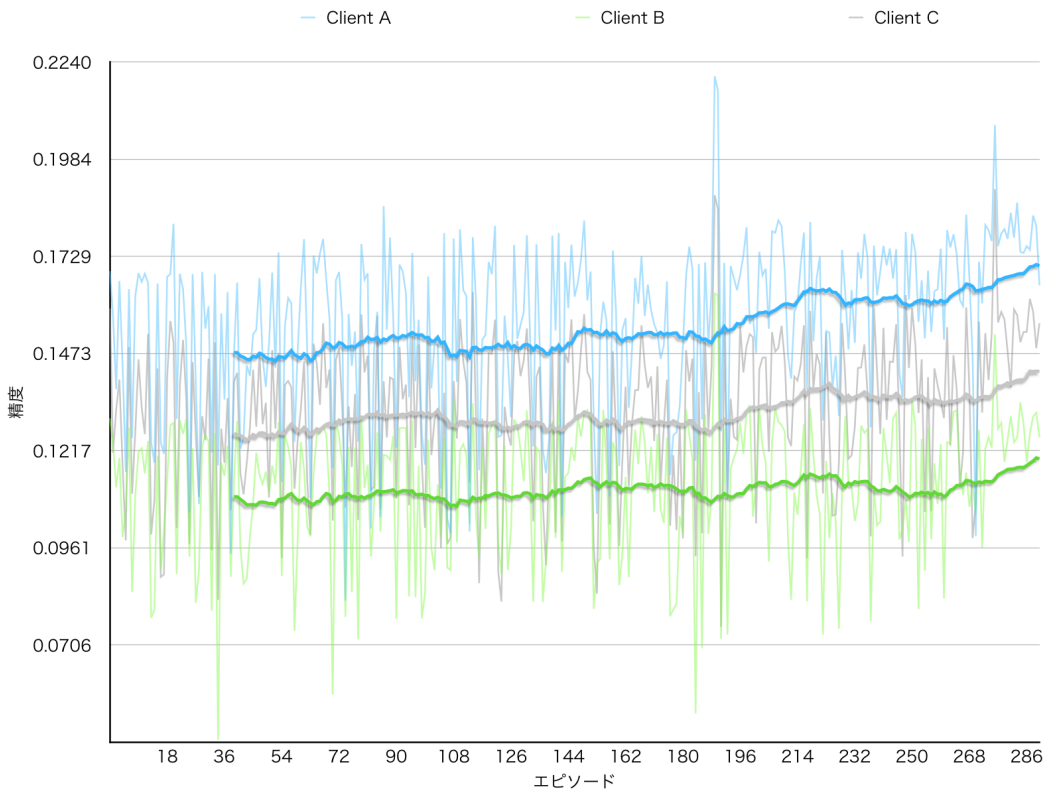

図 21: 強化学習の各エピソードで得たニューラル機械翻訳モデルの精度

表 12: 強化学習で得た最善な組織パターン

	1 回目	2 回目	3 回目	4 回目	5 回目
Client A	AC	AC	AC	AC	ABC
Client B	AC	AC	AC	C	AC
Client C	AC	BC	AC	AC	AC

織化で構築したニューラル機械翻訳モデルの精度の考察について述べる．

まず，サーバ最適の場合に，グリーディ法による自己組織化手法で構築したニューラル機械翻訳モデルの翻訳精度は自分野にも全分野にも従来手法で構築したニューラル機械翻訳モデルの翻訳精度より高いことが実験で証明された．よって，この提案手法で構築されたニューラル機械翻訳モデルの専門性，汎用性が従来より優れっていると言える．

各クライアントにフォーカスすると，自分野において，クライアント A の精度の上昇の幅が 43.28% 平均より上回る．それに対して，クライアント B の精

表 13: 最善な組織パターンで構築したニューラル機械翻訳モデルの評価結果

	学習データ	測定データ	評価データ	精度
Client A	伝統文化	伝統文化	伝統文化	0.22021
Client B	歴史	歴史	歴史	0.16306
Client C	人名	人名	人名	0.18891

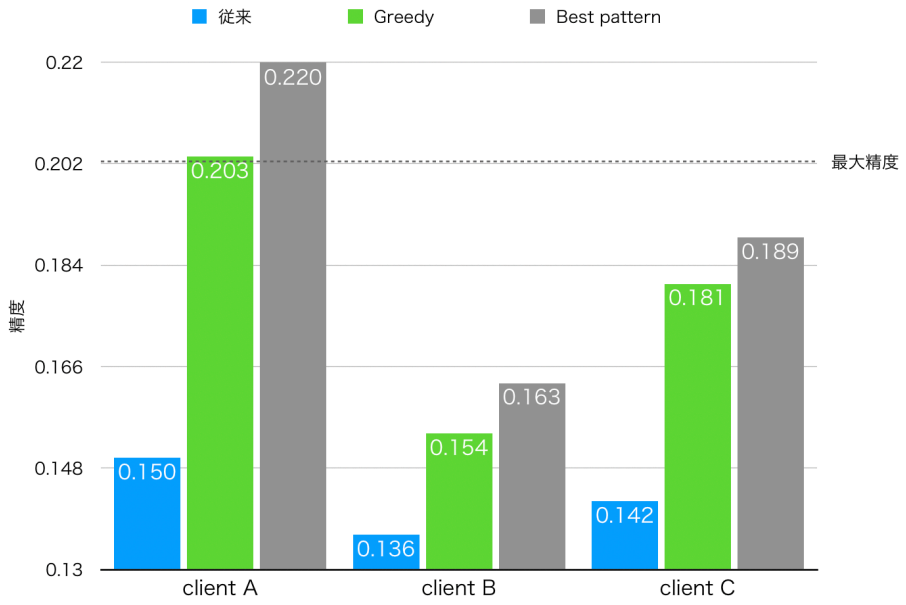

図 22: 最善な組織パターンとクライアント最適のグリーディ法と従来手法の比較

度の上昇の幅は三者の中に最下位 15.69% であり，クライアント A 精度の上昇量 4 割にいかない程度である．同じ精度が上昇したとはいえ，上昇量の差が激しいと感じられる．その原因は表 6 が示したように 5 回の集約の中に，クライアント B が組織から排除された回数が 3 回である．構築されたニューラル機械翻訳モデルの中，クライアント B が持つデータセットの特徴量がクライアント A とクライアント C より少ない．そのため，構築されたニューラル機械翻訳モデルはクライアント B が持つデータセットの分野（歴史）での表現がイマイチである．同様の理由で，5 回の集約の内にクライアン A が持つデータの特徴量が 4 回の集約に使われたため，構築されたニューラル機械翻訳モデルはクライアント A が持つデータセットの分野（伝統文化）でより良い表現ができる．

図 13 からサーバ最適で 5 回集約が完了時点で得た共通のニューラル機械翻訳モデルの精度が 0.19003 に対して，クライアントたちが自分が保有するデー

タで共通のモデルを仕上げし，得たモデルの精度はクライアント A が 0.01206，クライアント B が 0.01851，クライアント C が 0.01676 ほど減少した．表 6 のクライアントの組織履歴から見て，ワースト 2 回集約の際にクライアント A, B, C が一つの組織に形成し，三者が持つデータの特徴量がニューラル機械翻訳モデルに取り込んだ．しかし，最後の仕上げによってうまく統合された他のデータセットの特徴量と自分のデータセットの特徴量に打ち消される．学習が安定になる前に，クライアントたちの自己組織を干渉するのが精度が落ちた原因と推測する．また，推測を検証するために，仕上げたモデルをさらに集約し（6 回目），グリーディ法で選択されたモデルが集約モデル ABC である．共通テストデータで評価して，精度は 0.19625 であり，前回より上昇した．今回の実験設定の制約下に，サーバ最適で 5 回目に集約される時にニューラル機械翻訳モデルの学習がまだ安定していないことが判明した．

次に，クライアント最適の場合に，グリーディ法による自己組織化手法で構築したニューラル機械翻訳モデルの翻訳精度も自分野にも全分野にも従来手法で構築したニューラル機械翻訳モデルの翻訳精度より高いことが分かった．この提案手法で構築されたニューラル機械翻訳モデルの専門性，汎用性が従来手法より優秀であることが言える．

しかしながら，サーバ最適と同様に，各クライアントにフォーカスすると，クライアント A とクライアント B の精度上昇量の差が大きく，クライアント A の方の精度上昇量が多い．表 9 に示したクライアントの自己組織履歴からクライアント A が 5 回の集約は全部自分が参加した組織を選択した．自分が保有しているデータセットが自分と他人に学習され，学習結果を積極的に取り込む．その上に，第 2 回の集約の時にクライアント C と組織を組んで他のデータセットの長所も獲得した．そのため，クライアント A が構築されたニューラル機械翻訳モデルは自分野（伝統文化）で良い精度を取れた．

それに対して，クライアント B は他のクライアントの学習結果を自分より多く選択した．自分野のデータの学習不十分はクライアント B が構築されたニューラル機械翻訳モデルが良い評価をとれなかった理由と考える．

表 9 のクライアントの自己組織履歴を見て，クライアントたちが 3 回目以後に形成された組織が同じであることを明白である．連邦学習の最後にクライアントが各自のデータを使って，モデルを仕上げるではなく，もう一回グリーディ法で組織を形成した方がより高い精度のニューラル機械翻訳モデルが得ら

れるかもしれないと推測する．推測を検証するために，５回目で得た集約モデルに対してグリーディ法でさらに集約を行い，得たニューラル機械翻訳モデルの自分野での精度はそれぞれ 0.20327（クライアント A），0.16080（クライアント B），0.18829（クライアント C）である．表 10 に示した結果より高い（クライアント B と C）または一致（クライアント A）．したがって，組織を形成が安定になる場合，組織で構築したニューラル機械翻訳モデルは組織から離脱し，単体で仕上げしたニューラル機械翻訳モデルより高い精度になるという結論が考えられる．

そして，マルチエージェント深層強化学習による自己組織化手法の考察について述べる．

深層強化学習を通じて，構築されたニューラル機械翻訳モデルの精度がグリーディ法より高い複数の組織パターンが探測され，表 13 に示すとおり，構築されたニューラル機械翻訳モデルの精度が最も高くなる最善な組織パターンを発見した．

各集約の時にニューラル機械翻訳モデルの精度が最大化するではなく，各行動の長期的な価値を考えた上で行動を決定するため，最善な組織パターンを見つける原因だと考える．

また，今回の実験の中に 191 回目のエピソードに最善な組織パターンを発見した．三つクライアントのため，各クライアントは毎回集約の時に取れる行動数は 7 種である．1 エピソードに 5 回の集約が行われる．つまり，$(7\times7\times7)^5 = 4{,}747{,}561{,}509{,}943$ 種の異なるエピソードが存在する．もし，総当たり法で最善な組織パターンを探索しようとしたら，仮に 1 エピソードに１時間かかるとしても，この最善な組織パターンを見つけるまで最悪の場合は約 54 万年が必要となり，非常に非現実なことである．総当たり法と比べ，マルチエージェント深層強化学習による自己組織化手法は約一週間で最善な組織パターンを発見したのため，提案手法は実現可能と言える．

表 12 に載せているクライアント B の組織化パターンに注目すると，面白いことを気づける．クライアント B にペアする深層強化学習エージェントは 5 回連邦学習の集約の中にクライアント B が参加している組織を一回も選択していない．連邦学習の段階に自分自身が学習した結果を選ぶより直接に取り込まない方が，自分野での翻訳精度が高いニューラル機械翻訳モデルを構築できる可能性もある．

図 22 からクライアント A が最善な組織パターンで構築したニューラル機械翻訳モデルは伝統文化での翻訳精度が最大精度にみなした中央集権型で構築したニューラル機械翻訳モデルの翻訳精度より高くなった．その理由は集約の際にクライアント A がよく他にクライアントに組織され，クライアント A が伝統文化のデータセットを学習した結果が多く取り上げられる．同時に クライアント B があまり組織に参加できなかったことが原因で，総データの中クライアント A が持つデータの割合が上げ，式 3 により中央集権型よりクライアント A のデータがより十分に学習されたと考えられる．

図 21 に描いている各エピソードに各クライアントが構築されたニューラル機械翻訳モデルの精度の移動平均線により，180 エピソードあたりに勾配がきつくなり，深層強化学習エージェントが試行錯誤の中でより精度が高いニューラル機械翻訳モデルを構築できるクライアントの組織法を見つめ始め，実験が一時停止されるまで組織法を改善し続けることが分かった．

しかしながら，今回の実験が一時停止されるまで深層強化学習エージェントたちの学習はまだ収束されてない．安定的に高精度のニューラル機械翻訳モデルを構築できる組織パターンを選択できるためにまだエピソードを増やし，深層強化学習に学習コストをかけ続ける必要がある．収束しづらい原因についていくつの点を考えられる．まず，深層強化学習エージェントたちの状態空間のサイズは 10,000 であり，状態空間の各次元は 0 と 1 の間に十五桁の小数で表現する．そのため，状態空間が非常に大きくのが原因の一つと考えられる．そして，今回の実験は状態空間が大きゆえ，マルチエージェント深層強化学習の学習方法は各エージェントが独立して学習，行動を決定する完全非中央集権型に採用した．図 8 に示したように，完全非中央集権型が必要な学習空間が比較的に少ない分，学習が不安定，収束しにくいことがもう一つの原因と取り上げられる．

測定データのサイズを減らし，状態空間の大きさを抑えた上に，学習方法がより安定的な完全中央集権型か混合型にすることで，少ない深層強化学習の学習コストで安定的に高精度のニューラル機械翻訳モデルを構築できる組織パターンを選択できる深層強化学習が得られると予測できる．

以上のことから，以下の結論が導ける．ニューラル機械翻訳における，連邦学習の中に動的に連携相手を選択する自己組織化手法手法は参加者を一括に統合する従来手法より専門性と汎用性が高いニューラル機械翻訳モデルを構築で

きる．しかし，マルチエージェント深層強化学習による自己組織化の際に，強化学習エージェントの学習コストのことを考えると，学習コストを抑え，汎用的なエージェントを得るため，改善の余地が存在する．

第7章　おわりに

本研究では，従来の連邦学習の参加者を一括に統合する手法の代わりに，参加者たちが動的に自己組織する手法を提案した．そして，グリーディ法による自己組織化手法は，従来手法である Federated Averaging より高精度のニューラル機械翻訳モデルを構築できることで提案手法の有効性を示した．そして，マルチエージェント深層強化学習を導入することによって，より適切な組織パターンを探索し，構築されたニューラル機械翻訳モデルの精度をさらに引き上げた．

本研究の貢献は以下の 2 点である．

動的組織法

動的組織法と Federated Averaging によって構築されたニューラル機械翻訳モデルを自ドメインと他ドメインの翻訳精度を評価し，動的組織法によって構築されたモデルが自ドメインの精度は 28.99%高くなり，他ドメインの精度にも 16.83%向上し，提案手法の有効性を検証した．

最善な組織パターンの探索

マルチエージェント深層強化学習の導入によって，連邦学習の参加者たちが将来的のモデルの精度上昇を考慮しながら，最適な連携相手を動的に選択することで自己組織化する手法を考案した．動的組織法によって構築されたニューラル機械翻訳モデルと比較し，6.33%精度を向上し，従来の中央集権型ニューラル機械翻訳の性能の 94.26%に達した．

本研究の検証は，連邦学習に三つの参加者だけが存在する小規模かつシンプルな環境で行われた．従来手法より動的に連携相手を選択する自己組織化手法手法の方が高精度のニューラル機械翻訳モデルを構築でき，提案手法の有効性を証明したが，参加者の数が増えると，参加者たちの組み合わせ数にも急激に増える．それに伴い，グリーディ法による自己組織化に集約時に組み合わせごとに集約モデルを生成するコスト，最適な集約モデルを選出コストが爆発することが予想できる．参加者が増えた時にコストを抑えるアルゴリズムを考案する必要がある．

また，マルチエージェント深層強化学習による自己組織化はどの参加者とどの参加者を組織になることを行動としてモデル化した．参加者の人員交代が発生された時に，多大なコストをかけ，新しい深層強化学習エージェントを訓練

しなければならない．連邦学習の構成が変わっても，対応できるマルチエージェント深層強化学習のモデルを考案する必要がある．

謝辞

本研究を行うにあたり, 研究設備を購入し，熱心なご指導, ご助言を賜りました指導教官の村上陽平准教授に深謝申し上げます．また，普段からご協力，ご意見を下さいました丁好さんにも感謝の意を表します.

参考文献

[1] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y. Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data," in *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics* (A. Singh and J. Zhu, eds.), vol. 54 of *Proceedings of Machine Learning Research*, pp. 1273–1282, PMLR, 20–22 Apr 2017.

[2] 中澤敏明, "機械翻訳の新しいパラダイム：ニューラル機械翻訳の原理," **情報管理**, vol. 60, no. 5, pp. 299–306, 2017.

[3] 清藤武暢, "プライバシー保護技術としての連合学習の仕組みと最新動向," **電子情報通信学会 基礎・境界ソサイエティ** *Fundamentals Review*, vol. 16, no. 3, pp. 196–204, 2023.

[4] 丁好, "連邦学習を用いた非中央集権型のニューラル機械翻訳の有用性検証.".

[5] T. Roosta, P. Passban, and A. Chadha, "Communication-efficient federated learning for neural machine translation," 2021.

[6] C. J. C. H. Watkins and P. Dayan, "Technical note q-learning.," *Mach. Learn.*, vol. 8, pp. 279–292, 1992.

[7] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, "Playing atari with deep reinforcement learning," 2013. cite arxiv:1312.5602Comment: NIPS Deep Learning Workshop 2013.

[8] H. van Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double q-learning," 2015. cite arxiv:1509.06461Comment: AAAI 2016.

[9] G. Klein, Y. Kim, Y. Deng, J. Senellart, and A. Rush, "OpenNMT: Open-source toolkit for neural machine translation," in *Proceedings of ACL 2017, System Demonstrations*, (Vancouver, Canada), pp. 67–72, Association for Computational Linguistics, July 2017.

[10] M. Mohri, G. Sivek, and A. T. Suresh, "Agnostic federated learning," in *Proceedings of the 36th International Conference on Machine Learning* (K. Chaudhuri and R. Salakhutdinov, eds.), vol. 97 of *Proceedings of Machine Learning Research*, pp. 4615–4625, PMLR, 09–15 Jun 2019.

[11] 米谷竜, "連合学習入門," **精密工学会誌**, vol. 87, no. 8, pp. 662–665, 2021.