

## 深層強化学習による営業活動意思決定支援システムの優位性分析と改善の考察 Superiority Analysis and Improvement Considerations of Decision-making Support System in Sales Activities by Deep Reinforcement Learning

中山 義人<sup>†</sup> 斎藤 忍<sup>‡</sup> 森 雅広<sup>†</sup>  
Yoshihito Nakayama Shinobu Saito Masahiro Mori

### 1. はじめに

近年の急激な環境変化により、企業の意味決定に関わる要因や影響が複雑化するとともに、意思決定のスピードも求められる時代になってきた[1]。あらかじめ定まったルールのない非定型業務における意思決定のシーンでは、膨大な可能性の中から有効な意思決定の候補を見つけることが重要な課題となっている[1]。この課題の解決には、機械学習によるアプローチが不可欠である。機械学習の適用により、これまで属人性の高かった意思決定を支援する仕組みが求められている。筆者らは、企業活動のなかでもより複雑性が高く、かつ非定型の業務領域である営業活動の意思決定に注目して研究を進めてきた[2],[3]。

営業活動では、営業担当者が顧客とのコミュニケーションを通じて営業知識を動員し、場面に応じて顧客に対する適切な営業のアクティビティ（例：ユーザの状況把握、案件の必然性確認、キーマン把握）を選択する（図1）。この意思決定を支援するシステム（意思決定支援システム）には、顧客との間でつねに変化する場面に応じた適切なアクティビティのリコメンドを、営業担当者に提示することが求められる。意思決定支援システムが、営業活動のプロセス全般を通じてリコメンドを繰り返すことで、営業担当者の活動を受注に誘導し、高い売上を実現することが可能となる。さらには、受注成果の高い営業担当者の意思決定が組織全員で再現可能となり、組織力強化にもつながる。

以上の考えに基づき、これまでに筆者らは、受注確率の高い営業担当者の営業知識が反映されたプロセスモデルの構築、およびそれに基づいたアクティビティのリコメンドを可能とする営業活動意思決定支援システムを構築している[4][5]。当該システムはイベントログからプロセスモデルを構築する技術であるプロセスマイニング[6],[7]を応用している。初めに、社内に蓄積された営業日報データから、各アクティビティの選択が受注結果に与える影響を受注確率として推定する。次に、受注確率の高い営業活動プロセスの規則性を機械学習により抽出し、プロセスモデルを構築した。

前論文[5]では、プロセスモデル構築アルゴリズムとして、部分観測マルコフ決定過程[8]（以下、POMDP）、および環境モデルとセルフプレイを用いた深層強化学習[9]の2つのアルゴリズムを適用した営業活動意思決定支援システムを構築した。更に両アルゴリズムによるプロセスモデルからリコメンドされたアクティビティの推移を、筆者の一部が所属する企業の法人営業グループの熟練営業担当者20名

に確認してもらった。その後、アンケートを実施してリコメンドの納得度を測定した。その結果、深層強化学習に対する評価（経験合致度）がPOMDPよりも高いことが説明できた。

一方で、両プロセスモデルがリコメンドするアクティビティの系列には、具体的にどのような違いが生み出され、且つ、その違いが熟練営業担当者の深層強化学習に対する高い納得感にどのように影響したのかまでの詳細な調査には至っていない。

そこで本稿では、両プロセスモデルが提示するリコメンドの内容の調査・比較分析を通じて、深層強化学習に基づくシステムの優位性の要因、つまり熟練営業担当者の納得度が高くなった要因を調査する。その上で、深層強化学習に基づいたリコメンドに対する納得度において、高評価・低評価の比較分析により、リコメンドするアクティビティの内容の改善・高度化への方向性を考察する。

本論文の構成は以下のとおりである。2章「背景」では、筆者らの既存研究（POMDPと深層強化学習によるプロセスモデル構築アルゴリズムの概要と評価結果[4][5]）を述べる。3章「評価戦略」では、両プロセスモデルが提示するリコメンドの内容の調査・比較の方向性について解説し、さらに4章「結果」では、その評価結果を示した。最後に5章「考察」、6章「まとめ」を述べる。

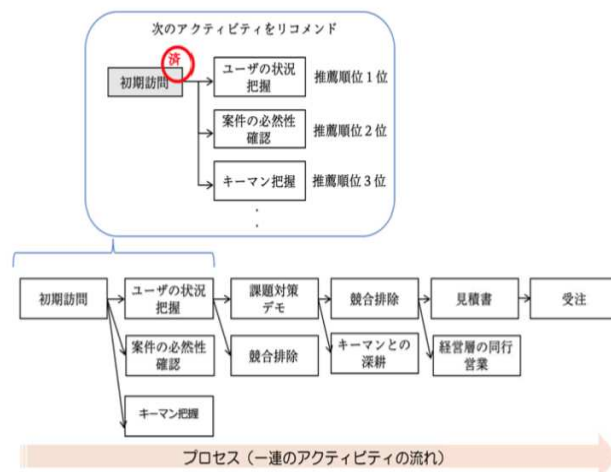


図1 営業活動意思決定支援システムのイメージ  
Fig. 1 Decision support system of sales activity.

<sup>†</sup> 株式会社 NTT データイントラマート  
NTT DATA INTRAMART Corporation  
<sup>‡</sup> 日本電信電話株式会社 NTT Corporation

2. 背景：POMDP と深層強化学習を適用した営業活動意思決定支援システム

2.1 案件シミュレーションによる教師データの作成

POMDP や深層強化学習では、前処理として大量の教師データを必要とする一方、実際の営業活動（日報メールなど）から取得できるデータには限りがある。そのため、営業をエージェント、顧客を環境と位置付けた案件シミュレータを開発し、ルールベースの営業日報データを大量に自動生成した[4]。

以降では案件シミュレータの詳細を述べる（図2）。営業パラメータ、顧客パラメータの全項目が正規分布で初期設定された状態から始まる。その後、営業状態に基づくルールベースのアクティビティ候補の中からランダムにアクティビティが選択され実施される。次に、顧客側では正規分布で重み付けされた顧客特性に従い顧客状態が更新される。その後、更新された顧客状態に即した顧客反応が後述するフィードバックルールから選択され、最後に、顧客反応に対応したルールベースの重み付けにより営業パラメータが更新される。

フィードバックルールは、熟練営業担当者により、計2046件の実際の営業活動から取得した日報データを用いて、アクティビティと顧客状態の区分ごとの顧客反応の組み合わせを設定した（図3）。

以上の案件シミュレータにより、案件データ総数 10,000 件と、営業日報データであるアクティビティデータ総数 267,000 件のデータセットが教師データとして作成した。

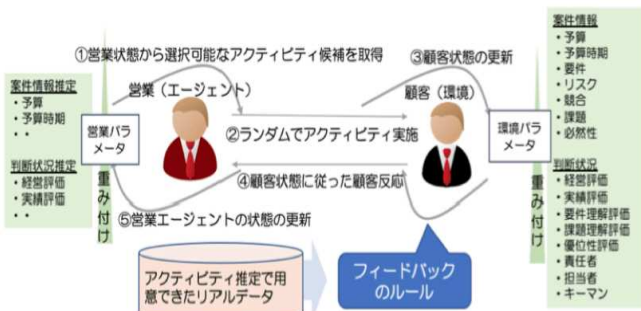


図2 案件シミュレータの詳細  
Fig. 2 Details of case simulator.

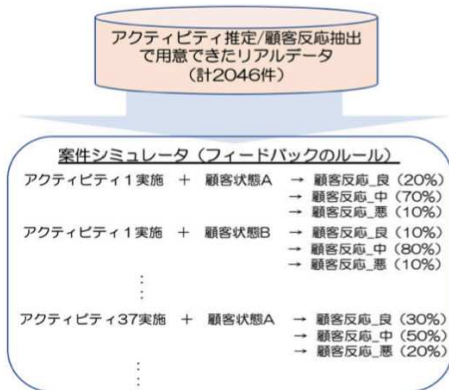


図3 フィードバックルールの詳細  
Fig. 3 Detail of feedback rules.

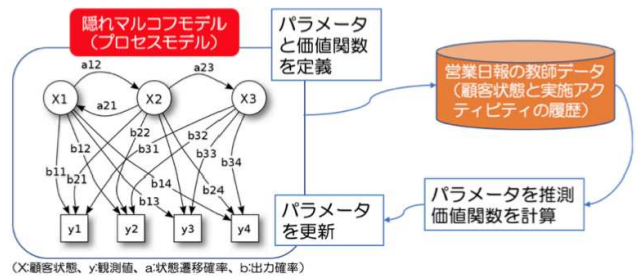


図4 隠れマルコフモデルによるプロセスモデル  
Fig. 4 Process model with hidden Markov model.

2.2 POMDP によるプロセスモデル構築の概要

前節の教師データを用いて POMDP によるプロセスモデルを構築した[4]。形式的には、POMDP は隠れマルコフモデルに行動、および行動を変更する動機を与える報酬を付与したものと解釈できる[10]。そこで、POMDP の近似手法として、隠れマルコフモデルによる状態観測に対して、方策学習による行動制御を組み合わせたプロセスモデルを構築した（図4）。

2.3 深層強化学習によるプロセスモデル構築の概要

前述の教師データの内容は、フィードバックルースに基づき設定した、顧客反応とアクティビティの組合せのセットである。そのため、前節の POMDP の学習モデルでは、もしフィードバックルールには存在しない想定外のアクティビティが提示された場合には、現実の営業活動と乖離した顧客反応を返してしまう可能性がある。そこで、上記の案件シミュレータにより生成される同一の教師データに深層学習を実施した。教師データに含まれる特徴を多段的により深く学習する。これにより、想定外のアクティビティにも現実に即した顧客反応の推定が可能となる。

深層強化学習によるアルゴリズムは 2 つのモデルを用いる（図5）。顧客状態をシミュレートした環境モデルと、営業担当者に受注確率の高いアクティビティをリコメンドするプロセスモデルである。環境モデルは、案件の背景、要件、顧客のパーソナリティといった顧客の環境を深層学習することで構築する。これにより、現実の顧客反応に即したモデルとなる。プロセスモデルは、環境モデルを用いた営業活動のシミュレーション（セルフプレイ[11]）を用いた強化学習により構築される。案件ごとの営業日報のデータセットから深層学習による環境モデルの構築を行い、最適なアクティビティ選択を環境モデルと営業エージェントによるセルフプレイ探索で学習する仕組みである[12]。

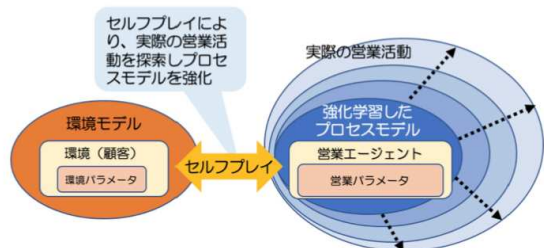


図5 深層強化学習によるプロセスモデルの構築  
Fig. 5 Process model with deep reinforcement learning.

## 2.4 プロセスモデルからリコメンドされるアクティビティの例

POMDP と深層強化学習によるプロセスモデルに基づいた営業活動意思決定支援システムからリコメンド可能なアクティビティは合計 38 種類ある。それぞれ営業活動の進展に合わせて D フェーズから A フェーズに分類されている。それぞれのフェーズごとに、一般的に理想とされるアクティビティの実施順序があり、表 1 では No で表現されている。法人向けの営業では骨格はほとんど同じであり、営業の一般的な教科書にも掲載されている[13]。筆者の一部が所属する企業の法人営業グループでも 10 年以上前より、このフェーズとアクティビティの実施順序に従って営業活動が管理されている。熟練営業担当者にもよく理解されている。

図 6 に、例として、POMDP によるプロセスモデルからリコメンドされたアクティビティの推移と、その中でもクロージング段階のリコメンドを拡大して可視化した。1 案件を横 1 列として、プロセスモデルからリコメンドされたアクティビティの種類が色の違いで可視化されている(左端から案件はスタートし右に向かってアクティビティが推移する)。

## 3. 評価戦略

営業活動意思決定支援システムがリコメンドするアクティビティの内容の改善・高度化に向けて、両プロセスモデルが提示するリコメンドの中身、特に「提案辞退」以外のリコメンドについて重点的に調査・分析を行い、考察を深めることとする。本稿で「提案辞退」以外を詳しく調べる理由として、前論文[5]では POMDP によるプロセスモデルからリコメンドされる「提案辞退」のアクティビティに営業担当者の関心が集約されてしまったことが挙げられる。彼らの多くは POMDP がリコメンドする「提案辞退」のアクティビティに違和感を持ち、結果として POMDP の評価が低下し、深層強化学習が優位となる一因となった[5]。一方、「提案辞退」のみに関心が集約された結果、POMDP がリコメンドする他のアクティビティに関する分析や評価の実施には至っていない。そこで本稿では、「提案辞退」以外のアクティビティに関して、深層強化学習に基づくシステムの優位性の要因、つまり熟練営業担当者の納得度が高くなった理由を調査することにする。さらに、深層強化学習に基づいたリコメンドに対する納得度において、高評価・低評価の比較分析により、リコメンドするアクティビティの内容の改善・高度化への方向性を考察する。以降ではそれぞれの評価戦略を述べる。

### 3.1 評価戦略 1 : POMDP と深層強化学習の受注案件のアクティビティ推移の違い

営業活動において受注を決定づけるのは、最終局面でのアクティビティ選択である[14][15]。この最も重要な局面は「クロージング」とも呼ばれており、最後に顧客が受注に踏み切ることができるように、購入の意思を確かめながら受注しやすい状況を作りだすことに努める事になる。顧客は基本的に、どの商品/サービスを選ぶのか悩んでいる状態なので、適切なクロージングを行なうことが顧客の検討状況を後押しすることにつながる。営業担当者はクロージングまでの顧客を誘導する意識を持つことが大切で

ある。このことはコロナ禍以降で急速に進展しているオンライン営業において一層重要性を増している。

そこで評価戦略 1 では、両プロセスモデルからリコメンドされる受注案件のアクティビティ推移の違いについて、クロージング段階のアクティビティを重点的に調査する。この調査を通じて、前論文における深層強化学習に基づくシステムの優位性の要因、つまり熟練営業担当者の納得度が高くなった理由を分析する。

表 1 法人向け営業活動のアクティビティ実施順序  
Table.1 Activity order of sales activities.

no	act	phase	no	act	phase
1	コネ把握	D	25	提案体制の決定	B
2	経営評価重要度	D	26	同行営業	B
3	実績評価重要度	D	27	製品説明	B
4	要件理解評価重要度	D	28	課題対応	B
5	課題理解評価重要度	D	29	競合排除	B
6	優位性評価重要度	D	30	深耕営業	B
7	責任者重要度	D	31	責任体制の決定	B
8	担当者重要度	D	32	受注判定会議	A
9	キーマン重要度	D	33	提案書作成依頼	A
10	初期訪問	D	34	提案書作成	A
11	要件取得	C	35	提案書プレゼンテーション	A
12	ユーザ状況把握	C	36	契約情報	A
13	課題取得	C	37	見積り	A
14	予定時期取得	C	38	再見積り	A
15	予算取得	C		受注	
16	体制取得	C		失注	
17	リスク確認	C		提案辞退	
18	競合取得	C			
19	ステークホルダー情報取得	C			
20	必然性確認	C			
21	提案方針の検討	C			
22	訴求ポイント	C			
23	戦略会議	C			
24	提案方針の決定	C			

D: 訪問  
C:ヒヤリング  
B:初期提案  
A:最終提案&クロージング



図 6 リコメンドされたアクティビティ推移  
Fig.6 Recommended activity transition

### 3.2 評価戦略 2 : 深層強化学習のアンケートにおける高評価と低評価の違い

POMDP に対して優位性が確認できている深層強化学習によるリコメンド (受注案件) においても, 熟練営業担当者による納得度の高い評価と低い評価のアクティビティ推移が存在していることが確認できている. そこで本評価戦略では, 深層強化学習によるアルゴリズムがリコメンドするアクティビティの内容の改善・高度化に向けた調査を行う. 具体的には, 高い評価と低い評価のアクティビティ推移の比較分析を通じて, 熟練営業担当者の納得度が異なる要因を調査する. 低い評価につながる要因がわかれば, それらの要因を排除した上でプロセスモデルを再作成することができる. これにより, リコメンドするアクティビティの内容の改善・高度化への方向性が明らかにできる. 一方, 熟練営業担当者の評価の高低は, クロージング段階だけで決まるものではない. 評価戦略 2 では営業活動のプロセス全般にわたって比較分析を行う.

### 3.3 評価用データの準備

前論文[5]では, 両プロセスモデルからリコメンドされたアクティビティ推移が, POMDP : 200 案件, 深層強化学習 : 200 案件の合計 400 案件用意され, 営業担当者 20 名にそれぞれ 10 案件ずつ計 20 案件がアンケート評価に用いられた.

評価戦略 1 ではこれらの案件から, 前処理として「提案辞退」および「失注」で終了する案件を除き, 「受注」で終了する案件のみを抽出した. POMDP では 200 案件中 124 案件, 深層強化学習では 200 案件中 116 案件が抽出された. その上で, 評価戦略 1 の評価データとして, そこからクロージングの局面として, 「受注」を含めた最終 10 アクティビティの推移のみを抽出した. 評価戦略 2 の評価データとして, 深層強化学習からリコメンドされた 200 案件を, 高評価 (合致する (=3)), 中評価 (おおむね合致 (=2)), 低評価 (少し合致 (=1), 合致しない (=0)) に分類したところ, 高評価 64 案件, 中評価 36 案件, 低評価 16 案件を準備することができた.

## 4. 結果

### 4.1 評価戦略 1 の結果

3.3 節の両プロセスモデルからリコメンドとして出力されたアクティビティの出現回数の差異を確認した. クロージングの局面となる受注を含めた最終 10 アクティビティの推移のみを対象としている.

表 2 は, 前章で準備した評価戦略 1 の評価データにおいて, リコメンドされたアクティビティを出現頻度の高い順番に掲載している. 併せて, アクティビティの所属するフェーズも記されている. また表 2 中のグラフはフェーズごとの出現頻度を表している. 出現頻度で 10%以上のアクティビティを黄色で強調している. 深層強化学習においては「課題対応(B) : 33%」「提案方針の検討(A) : 23%」「提案書プレゼンテーション(A) : 21%」の 3 つのアクティビティだけで合計 77%を占めている. 一方で, POMDP では最も多い「深耕営業(B) : 19%」を筆頭に, 「ステークホルダー把握(C) : 17%」「競合排除(B) : 14%」「訴求ポイント

(C) : 11%」「同行営業(B) : 10%」となり, 出現しているアクティビティの内容と出現頻度に差がある.

つまり, 深層強化学習では A フェーズのクロージング段階で出現する上位のアクティビティが集中的にリコメンドされているのに対して, POMDP では A フェーズは少なく C フェーズが多く出現している. 営業担当者にとって最も重要な局面である「クロージング」において, 深層強化学習ではその段階で実施すべき A フェーズのアクティビティが適切にリコメンドされている. このことが, 熟練営業担当者の納得度の高さにつながっていると推測される.

次に, 一般的なプロセスマイニングツール[16]に具備されているプロセス可視化機能を利用して, リコメンドされるアクティビティの遷移フローについて調査した. 本論文では, オープンソースのプロセスマイニングツールとして利用の多い ProM (Process Mining Framework) [17][18]を採用し, 3.3 節の評価戦略 1 用の評価データを ProM に取り込むことにより, その遷移フローをビジュアルに可視化した.

POMDP のプロセスモデルによるアクティビティの遷移を可視化したものが図 7 上側の遷移フロー, 深層強化学習のプロセスモデルによるアクティビティの遷移を可視化したものが図 8 上側の遷移フローである. また, ProM には, 遷移フローの中で, 通過率が高いアクティビティを抽出する機能がある (ノードの重要度フィルター[19]). そこで, 通過率 20%以下のアクティビティを足切りすることで, 遷移フローの特徴を際立たせてビジュアル化した (図 7, 図 8 の下側の遷移フロー). それぞれの図中で通過率 20%以下のアクティビティは青い八角形に集約されている. 例えば図 7 では, 上側の遷移フロー中で点線赤丸で囲ったアクティビティは, それぞれ通過率 20%以下であったので, 下側の遷移フローでは集約して「Cluster26」となっている.

また, 遷移フローの中でも, 全体の中央値以上の出現確率を持つアクティビティ間の遷移フローを太線で, それ以外を細線で表現することで, 大まかな遷移フローの流れを把握しやすくなった. 出現確率とは, 全ての遷移フローに占める特定アクティビティ間の遷移フロー数の割合である.

表 2 アクティビティ出現回数とフェーズごと出現頻度  
Table.2 Number of activity occurrences and frequency of appearance for each phase.

POMDP			深層強化学習		
アクティビティ	出現頻度	フェーズ	アクティビティ	出現頻度	フェーズ
訴求ポイント	11%	C	課題対応	33%	B
深耕営業	19%	B	提案方針の検討	23%	A
ステークホルダー	18%	C	提案書プレゼン	21%	A
競合排除	15%	B	責任体制の決定	8%	B
同行営業	10%	B	製品説明	4%	B
課題対応	6%	B	提案書作成	3%	A
責任体制の決定	5%	B	競合排除	2%	C
提案書作成	5%	A	体制取得	1%	C
競合取得	4%	C	予定時期取得	1%	C
戦略会議	3%	C	深耕営業	1%	B
契約情報	1%	A	要件取得	1%	C
製品説明	1%	B	リスク確認	1%	C



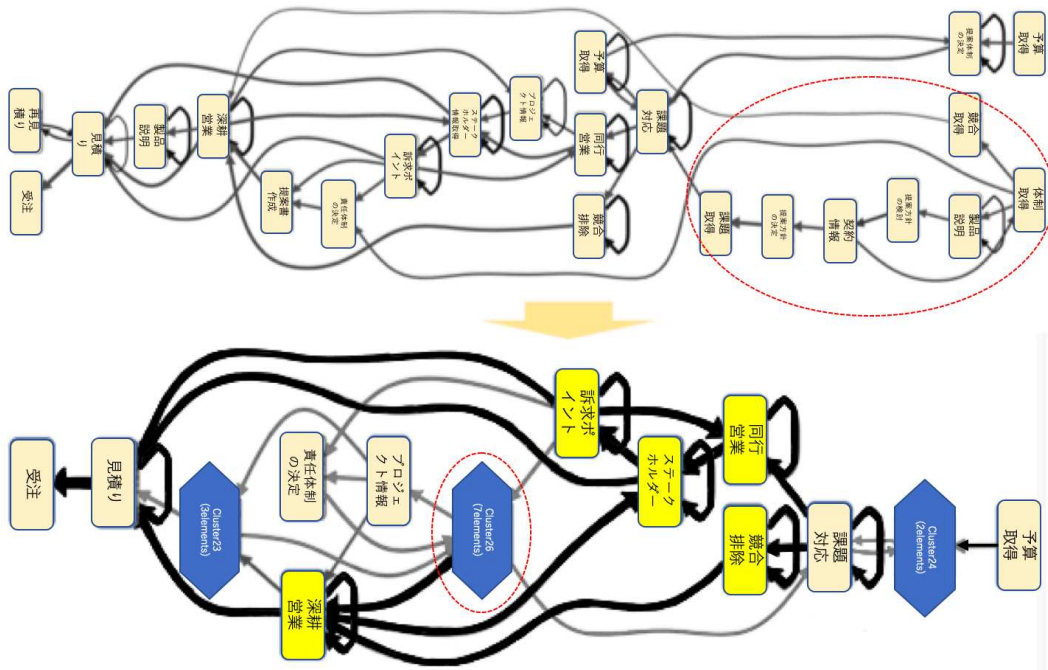


図 7 POMDP で構築したプロセスモデルよりリコメンドされたアクティビティの遷移フロー  
 Fig.7 Activity transition flow recommended by process model of POMDP

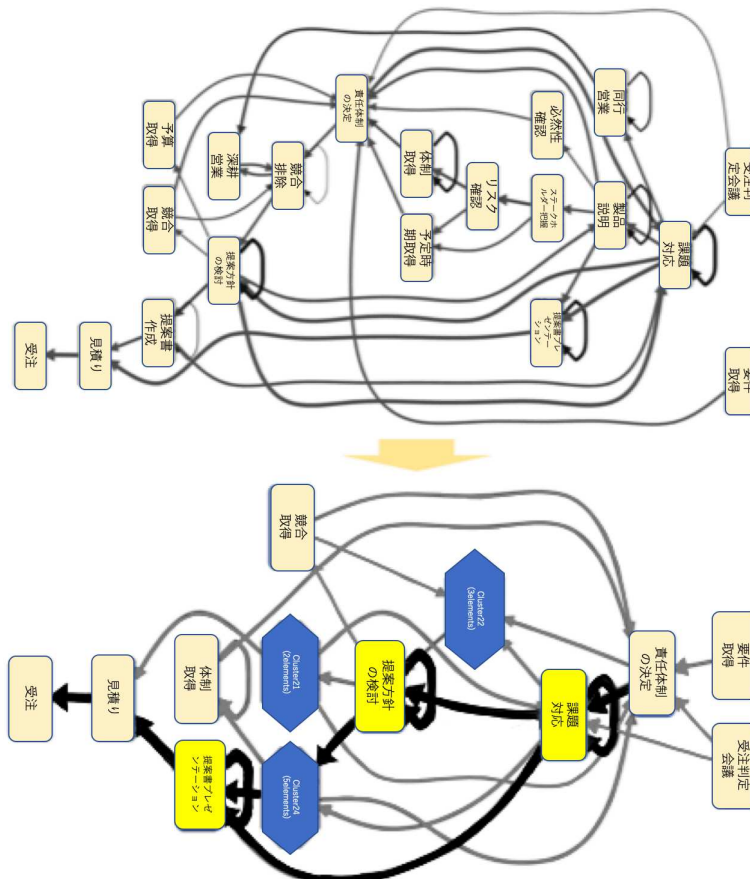


図 8 深層強化学習で構築したプロセスモデルよりリコメンドされたアクティビティの遷移フロー  
 Fig.8 Activity transition flow recommended by process model of DL

図 7 の POMDP のプロセスモデルによるアクティビティの遷移フローを、太線の遷移フローに着目して確認すると、表 2 に掲載した出現頻度 10%以上の 5 つの黄色いアクティビティ間のやりとりを経て、「見積り」「受注」のアクティビティに推移している。特に、「競合排除(B)」→「深耕営業(B)」→「見積り」につながる遷移フローと、「同行営業(B)」から「ステークホルダー把握(C)」「訴求ポイント(C)」「深耕営業(B)」のいずれかのアクティビティを経て「見積り」につながる遷移フローの遷移確率が高く特徴づけられた。また、太線の遷移フローの矢印の方向を確認すると、「同行営業(B)」→「ステークホルダー把握(C)」, および「深耕営業(B)」→「ステークホルダー把握(C)」へと、フェーズの進行方向が逆に戻る遷移確率も高く特徴づけられた。

図 8 の深層強化学習によるプロセスモデルでは、太線の遷移フローを確認すると、同じく表に掲載した出現頻度 10%以上の 3 つの黄色いアクティビティ間のやりとりを経て、「見積り」「受注」のアクティビティに推移している。

特に、「課題対応(B)」→「提案書プレゼンテーション(A)」→「見積り」につながる遷移フローと、「課題対応(B)」→「提案方針の検討(A)」→「通過率 20%以下のアクティビティ (青い八角系)」→「提案書プレゼンテーション(A)」→「見積り」につながる遷移フローの遷移確率が高く特徴づけられた。

以上より、深層強化学習によるプロセスモデルでは、B フェーズから A フェーズへと、クロージング段階で出現する上位のアクティビティを経て見積りまで着実に進行している。一方、POMDP によるプロセスモデルでは、A フェーズのアクティビティを経ずに見積りまで進行していること、さらには B フェーズから C フェーズへと進行方向が逆に戻る遷移フローが確認できた。

#### 4.2 評価戦略 2 の結果

表 3 では、納得度の高・中・低の評価ごとに、リコメンドされたアクティビティを出現頻度の高い順番に掲載している。アクティビティの所属するフェーズも併記している。出現頻度で 10%以上のアクティビティを黄色で強調したが、高・中・低評価の間で出現するアクティビティの内容に差は認められなかった。

図 9 はフェーズごとの出現頻度をグラフ表示している。C フェーズのアクティビティは、高評価<中評価<低評価の出現頻度となり、B フェーズのアクティビティは高評価>中評価>低評価の出現頻度、さらに A フェーズでは高評価>中評価>低評価の出現頻度であった。このことから、特に低評価では C フェーズのアクティビティに偏ったリコメンドがされているのに対して、中評価や高評価では A フェーズや B フェーズのアクティビティがバランスよくリコメンドされている。

次に、リコメンドされるアクティビティの遷移フローについて調査した。具体的には、高評価・中評価・低評価、それぞれにリコメンドされるアクティビティの出現順序が、D フェーズから A フェーズにかけて記載したアクティビティの理想的な実施順序 (表 2 の No) とどこまで適合・乖離しているかを調べた。その結果を以下の図 10 に示す。アクティビティの理想的な実施順序が  $y=x$  (図中の補助線) で示されるため、それとの乖離度を算出した。

表 3 評価結果に基づいたアクティビティ出現回数  
Table.3 Number of activity occurrences based on evaluation results.

高評価			中評価			低評価		
アクティビティ	出現頻度	フェーズ	アクティビティ	出現頻度	フェーズ	アクティビティ	出現頻度	フェーズ
要件取得	27%	C	要件取得	30%	C	要件取得	48%	C
課題対応	25%	B	課題対応	23%	B	課題対応	15%	B
提案方針の検討	16%	C	提案方針の検討	19%	C	提案方針の検討	16%	C
提案書プレゼン	12%	A	提案書プレゼン	8%	A	提案書プレゼン	8%	A
責任体制の決定	4%	B	責任体制の決定	3%	B	責任体制の決定	3%	B
競合排除	2%	B	製品説明	3%	B	リスク確認	3%	C
製品説明	2%	B	提案書作成	2%	A	製品説明	2%	B
同行営業	2%	B	ステークホルダー	2%	C	必然性確認	1%	C
体制取得	1%	C	体制取得	1%	C	深耕営業	1%	B
予定時期取得	1%	C	予定時期取得	1%	C	競合排除	1%	B
提案書作成	2%	A	必然性確認	1%	C	予算取得	1%	C
必然性確認	1%	C	深耕営業	1%	B	競合取得	1%	C
深耕営業	1%	B	競合排除	1%	B	体制取得	0%	C
ステークホルダー	1%	C	リスク確認	1%	C	予定時期取得	0%	C
リスク確認	1%	C	予算取得	1%	C	提案書作成	0%	A
予算取得	1%	C	同行営業	1%	B	ステークホルダー	0%	C
競合取得	1%	C	受注判定会議	0%	A	受注判定会議	0%	A
受注判定会議	0%	A	競合取得	0%	C	同行営業	0%	B
提案方針の決定	0%	C	提案方針の決定	0%	C	提案方針の決定	0%	C

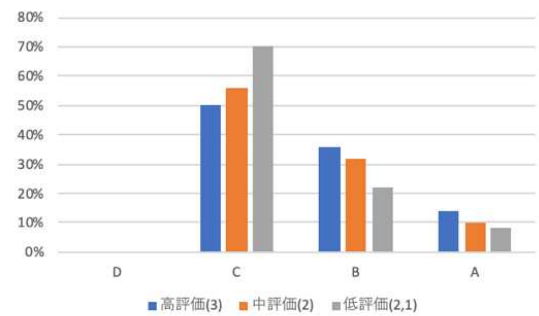


図 9 評価結果に基づいたフェーズごとの出現頻度  
Fig.9 Frequency of appearance for each phase based on evaluation results.

同じ現象を表す 2 つの観測値間の誤差の尺度である平均絶対誤差「Mean absolute error」と、さらに 2 つの観測値間の差の二乗を計算した上でその総和を取りデータの総数で割った値である平均二乗誤差「Mean squared error」による乖離度の測定を実施した。どちらも値が少ないほど誤差が少ないと言える。

その結果、中評価の乖離度が一番少なく、次に高評価、低評価の順であった。中評価と高評価の比較では期待と異なり、中評価の乖離度が少ない結果となった。低評価におけるアクティビティの出現順序は、中評価・高評価と比べて、理想とされるアクティビティの実施順序と乖離が大きく不規則であった。



横軸：アクティビティの実施順序(表1のNo.)  
縦軸：アクティビティの遷移フロー中の出現位置の平均

図 10 アクティビティの理想的な実施順序との乖離  
Fig.10 Deviation from the ideal order of sales activity.

## 5. 考察

### 5.1 評価結果の考察

評価戦略1では、POMDPと深層強化学習によるプロセスモデルのアクティビティ推移の違いについて、クロージ

ング段階のアクティビティを調査した。その結果、以下2点が明らかになった。

- ・深層強化学習では、Aフェーズのクロージング段階で出現する上位のアクティビティが集中的にリコメンドされているのに対して、POMDPではAフェーズは少なくCフェーズが多く出現している。

- ・深層強化学習では、BフェーズからAフェーズへと、クロージング段階で出現する上位のアクティビティを経て見送りまで着実に進行しているのに対して、POMDPではAフェーズのアクティビティを経ずに見送りまで進行している。さらにはBフェーズからCフェーズへと戻りの進行も多く出現している。

このことから、深層強化学習によるプロセスモデルに基づいたリコメンドは、フェーズを踏みながらAフェーズまで進展しているため、特に営業活動で重要なクロージングの局面で実施すべきことを適切にリコメンドできていることが、熟練営業担当者の納得度の高さにつながっていると推測される。さらには、POMDPのプロセスモデルに見られる進行フェーズが戻る遷移フローは、営業活動の進展に合わせてDフェーズからAフェーズに向かってアクティビティが実施されることに慣れている熟練営業担当者からすると、評価を下げる要因になったことが推測される。これは前論文の「提案辞退」のみを抽出した調査・分析では把握できなかった要因であり、深層強化学習に基づくプロセスモデルからのリコメンドの優位性をあらためて裏付けるものとなった。

評価戦略2では、リコメンドするアクティビティの内容のさらなる改善・高度化を目的に、深層強化学習に基づくプロセスモデルから提示されるリコメンドにおいて、熟練営業担当者の納得度の高い評価と低い評価のアクティビティ推移を比較分析した。その結果、以下の2点が確認できた。

- ・低評価ではCフェーズのアクティビティにリコメンドが集中しているのに対して、中評価・高評価ではAフェーズ、Bフェーズ、Cフェーズとそれぞれバランスよくリコメンドされている。

- ・理想とされるアクティビティの実施順序との適合性において、中評価・高評価のアクティビティ出現順序は低評価よりも高い。

この結果から、高・中評価のアクティビティの出現頻度は、低評価に比べて、フェーズごとにバランスよくリコメンドされていること、さらには高・中評価のアクティビティ実施順序は、低評価に比べて、理想的なアクティビティの実施順序との適合度が高いことが、熟練営業担当者による評価の高さにつながっていることが推測できる。熟練営業担当者は10年以上前より、表2のフェーズとアクティビティの実施順序に従って営業活動が管理されていることから、これと親和性の高いアクティビティ推移に、高い評価を与えたことは納得できる。

## 5.2 深層強化学習のプロセスモデル改善に向けた考察

営業活動意思決定支援システムのリコメンド精度をさらに向上するための展望について述べる。深層強化学習によるプロセスモデルのアップデートの方向性として、熟練営業担当者から低評価となるアクティビティの出現パターンを抑えつけることが考えられる。

具体的には、2.2.1 節で述べた案件シミュレータによる教師データのセットをクラスタリングして、低評価の特徴となっているクラス（フェーズのバランスが悪い案件、アクティビティ出現順序が理想と大きく乖離する案件）を削除し、プロセスモデルを再学習するという流れになる。この流れは、深層学習による環境モデルの作成時に、人間参加型深層学習 (human-in-the-loop) [20] を取り入れることであり、より人間と機械知能を組み合わせることで効果的な機械学習アルゴリズムを生成することが可能となる。結果として、熟練営業担当者による評価で、これまで以上に高評価を取れるプロセスモデルが構築できることになる。

## 6. まとめ

これまで、営業日報などの非構造化データを入力情報として、あらかじめ明確なルールが存在しない非定型プロセスから、受注確率の高い営業プロセスの規則性を抽出・構築したプロセスモデルに基づき、営業意思決定支援システムを提案し評価を行ってきた。

特に、本論文では、プロセスモデル構築アルゴリズムとして採用した POMDP と深層強化学習のそれぞれにおいて、提示されたリコメンドの中身の調査・比較分析を通じて、深層強化学習に基づくプロセスモデルのリコメンドの優位性の要因を、前論文[5]から引き続き調査した。さらに、深層強化学習に基づいたプロセスモデルのリコメンドに対する高評価・低評価の比較分析により、リコメンド精度のさらなる改善・高度化への方向性を考察した。

これらの意思決定支援システムの構築から評価、改善に至る一連の手法は、営業活動以外においても、複雑性が高くかつ非定型の業務領域の意思決定の効率化に寄与し、ひいては従業員の働き方を変革することにつながる提案になると考えている。

### 参考文献

- [1] 戦略プロポーザル：複雑社会における意思決定・合意形成を支える情報科学技術，国立研究開発法人科学技術振興機構研究開発戦略センター，CRDS-FY2017-SP-03 (2018)。
- [2] 中山義人，森 雅広，成末義哲，森川博之，“営業活動における意思決定のプロセス発見手法 プロセスマイニングの応用アプローチ”，情報システム学会誌，Vol. 14, No. 1, pp. 26-38, DOI: 10.19014/jissj.14.1.26 (2018)。
- [3] Nakayama, Y., Mori, M., Narusue, Y. and Morikawa, H. “The process discovery Approaches for decision making in sales activities”, Proc. SCIS&ISIS 2018, Okayama, Japan, pp.1394-1399, DOI: 10.1109/SCIS-ISIS.2018.00217 (2018)
- [4] 中山義人，森 雅広，斎藤 忍，成末義哲，森川博之，“非定型業務における意思決定支援システムの適用ステップの提案と実践”，情報処理学会論文誌トランザクションデジタルプラクティス，Vol. 2, No. 4, pp. 1-12 (2021)。
- [5] 中山義人，森 雅広，斎藤 忍，成末義哲，森川博之，“深層強化学習による営業活動意思決定支援システム”，情報処理学会論文誌，Vol. 63, No. 4, pp. 1008-1018 (2022)。
- [6] Van Der Aalst, W. “Process Mining : Discovery, Conformance and Enhancement of Business Processes”, Springer, Heidelberg (2011)。
- [7] 飯島 正，田端啓一，斎藤 忍，“プロセスマイニング・サーベイ (第 01 回：概要と基本概念)”，情報システム学会誌，Vol. 11, No. 2, pp. 20-53, DOI: 10.19014/jissj.11.2.20 (2017)。
- [8] 澁谷長史，“《第 4 回》部分観測マルコフ決定過程と強化学習”，計測と制御，Vol. 52, No. 4, pp. 374-380 (2013)。
- [9] Racaniere, S., Reichert, D., Weber, T., Wierstra, V.D., Buesing, L., Battaglia, P., Pascanu, R., Li, Y., Heess, N., et al., “Imagination-augmented agents for deep reinforcement learning”, Advances in Neural Information Processing Systems, Vol. 30, pp. 5692-5699 (2017)。
- [10] 木村 元，“部分観測マルコフ決定過程下での強化学習 (特集) 強化学習”，人工知能学会誌，Vol. 12, No. 6, pp. 822-830 (1997)。
- [11] Heinrich, J. and Silver, D. “Deep reinforcement learning from self-play in imperfect-information games”, arXiv preprint arXiv:1603.01121 (2016)。
- [12] 中山義人，森 雅広，成末義哲，森川博之，“営業活動の意思決定プロセス強化における環境モデルに基づくアプローチ”，情報処理学会研究報告グループウェアとネットワークサービス，Vol. 2019-GN-107, pp. 1-5 (July 2019)。
- [13] Brennan, R., Canning, L., and McDowell, R., “Business-to-business marketing”, SAGE Publications Limited (2020)。
- [14] Johlke, M.C. “Sales presentation skills and salesperson job performance”, Journal of Business & Industrial Marketing, Vol. 21, No. 5, pp. 311-319 (2006)。
- [15] Muir, J. “The perfect close: The Secret to closing sales”, Herriman, UT: Best Practice International (2016)。
- [16] 飯島 正，田端啓一，斎藤 忍，“プロセスマイニング・サーベイ (第 02 回：ツール)”，情報システム学会誌，Vol. 12, No. 1, pp. 1-25, DOI: 10.19014/jissj.12.1.1 (2016)。
- [17] Dongen, B.F., Medeiros, A.K.A., Verbeek, H.M.W., Weijters, A.J.M.M., and Van Der Aalst, W., “The prom framework: A new era in process mining tool support”, Applications and Theory of Petri Nets 2005: 26th ICATPN 2005, (eds.) G. Ciardo, and P. Darondeau, LNCS-3536, pp. 444-454, Springer (2005)。
- [18] Van Der Aalst, W., Van Dongen, B.F., Günther, C.W., Mans, R.S., Medeiros, A.K., Rozinat, A., et al., “Prom 4.0: Comprehensive support for real process analysis”, Petri Nets and Other Models of Concurrency : 28th ICATPN 2007, (eds.) J. Kleijn, and A. Yakovlev, LNCS-4546, pp. 484-494, Springer (2007)。
- [19] Günther, C.W. and Van Der Aalst, W., “Fuzzy mining-adaptive process simplification based on multi-perspective metrics, International conference on business process management”, pp. 328-343, Springer, Berlin, Heidelberg (2007)。
- [20] Karwowski, W., et al., “International encyclopedia of ergonomics and human factors-3 Volume Set”, CRC Press (2006)。