

特別研究報告

題目

能動推論に基づく搬送と倉庫内の状況監視を両立する
搬送ロボット制御手法の提案と評価

指導教員

村田 正幸 教授

報告者

福山 敦也

2024 年 (令和 5 年) 2 月 7 日

大阪大学 基礎工学部 情報科学科

内容梗概

近年、自律走行を行うことができるロボットが多く開発され、工場や物流倉庫内で多く活用されている。これらの工場や物流倉庫内で動く搬送ロボットには、作業者が積み下ろした荷物や運ばれてきた荷物を運搬する役割がある。従来の搬送ロボットでは、ある箇所の荷物を運び終えた時に次に向かうべき荷物の場所をロボット自身が把握できていない場合が多く、その場合ロボットはまず運ぶべき荷物の探索を行い、荷物の場所に到着してから目的地までの移動について考え始める必要がある。これを解決するためには、作業者の作業位置や、次に運搬すべき荷物の場所を早くに把握しておく必要がある。すなわち、目的地への移動だけでなく作業者の作業位置や、次に運搬すべき荷物の場所などの周囲の状況の把握のための行動が必要となり、状況を把握するための移動という行動と、障害物を回避しながら目的地へと到達するための移動という行動の両立が求められることとなる。

本報告では、荷物の搬送と倉庫内の作業や搬送すべき荷物の位置などの状況の把握を両立できる自律型搬送ロボットの制御手法を提案する。本手法ではこの制御を能動推論を用いて行う。能動推論では、情報が無い状態から観測と行動を繰り返し、自由エネルギーを最小化する状態に移るべき状態とし行動することを人の脳の行動原理とし、これに基づいて状態の推論を行う。物流倉庫内で動くロボットは、壁や棚、倉庫内で働く作業員などの障害物の回避と、荷物の運搬のため荷物の置かれている場所と倉庫の出口への往復を同時に実現する必要がある。これは状況が不明であるという情報が無い状態から、探索と状況の観測を繰り返し目的地という望ましい場所へと移動を行うという解釈ができ、これは能動推論が考える人の脳の行動原理と一致する。そして、状況観測と目的地への移動という内容は、能動推論において統合的に扱うことができる。そこで本報告では、搬送ロボットにおける荷物の搬送先である目的地へと移動する行動と、自身の周囲を観測することにより倉庫内の状況を推測する過程をモデル化することで、搬送ロボットの制御を能動推論により行う手法を提案する。

能動推論での対象のエージェントをロボットに、観測を作業員や障害物の有無、状態をロボットの位置や倉庫の状況、ポリシーを移動として生成モデルを定義し、生成モデルで尤度、状態の遷移確

率，選好，初期状態，ポリシーの事前信念を表す5つの変数行列を対象の倉庫内の作業者などの障害物や目的地の位置から設定し，能動推論を用いて行動を決定する手法を実装し，能動推論を用いずに行動決定を行うロボットと比較し評価を行った．シミュレーションの結果，能動推論を用いることで単に目的地へと向かう行動だけでなく，同時に倉庫内の作業者や荷物の位置といった状況を把握するための行動もし，ロボットが荷物の運搬途中で次に運ぶべき荷物の位置や障害物の位置などの倉庫の状況を把握することが可能であることを示した．また，パラメータを調整することで運搬効率に大きな影響を与える場合は状況把握を行わないようにすることも可能であることを示した．

主な用語

自由エネルギー原理

能動推論

搬送ロボット

自律制御

経路計画

観測

目次

1	はじめに	6
2	能動推論	8
2.1	能動推論の概要	8
2.2	能動推論における生成モデルの設計	11
2.3	能動推論を応用した従来手法	13
3	能動推論に基づく搬送と倉庫内の状況監視を両立する	
	搬送ロボット制御手法	14
3.1	提案の概要	14
3.2	提案手法における生成モデルの詳細	15
4	提案手法の評価	19
4.1	設定	19
4.1.1	環境	19
4.1.2	比較対象	21
4.1.3	評価指標	22
4.2	評価結果	22
4.3	考察	29
5	おわりに	32
	謝辞	33
	参考文献	34

図目次

1	生成プロセスと生成モデル, 文献 [1] より引用, p23 図 2.2	8
2	生成モデルの構造, 文献 [1] より引用, p70 図 4.3	9
3	設計した生成モデル	14
4	倉庫の内部構造	19
5	作業者が移動しない場合の作業人や荷物の位置の設定	20
6	作業者が移動する場合の作業人や荷物の位置の設定	20
7	図 5(a) でロボットが進んだ経路	23
8	図 5(b) でロボットが進んだ経路	23
9	図 5(c) でロボットが進んだ経路	24
10	図 5(d) でロボットが進んだ経路	24
11	図 5(e) でロボットが進んだ経路	25
12	作業者が移動しない場合での作業人の位置に関する状態の信念の変化 (図 5(a)) . . .	25
13	作業者が移動しない場合での作業人の位置に関する状態の信念の変化 (図 5(b)) . . .	26
14	作業者が移動しない場合での作業人の位置に関する状態の信念の変化 (図 5(c)) . . .	26
15	作業者が移動しない場合での作業人の位置に関する状態の信念の変化 (図 5(d)) . . .	26
16	作業者が移動しない場合での作業人の位置に関する状態の信念の変化 (図 5(e)) . . .	27
17	作業者が移動する場合での作業人の位置に関する状態の信念の変化 (図 6)	28

表 目 次

1	ロボットが移動した経路長の合計	29
2	能動推論を用いて行動決定を行うロボットの経路長の増加率	29
3	能動推論を用いて行動決定を行うロボットの状況把握にかかる時間 (ステップ数)	30

1 はじめに

近年、自律走行を行うことができるロボットが多く開発されている。中でも工場や物流倉庫内などの施設において、荷物の運搬などを自動で行う搬送ロボットが多く導入され、普及している。これらのロボットには磁気テープや電磁誘導ケーブル等がひかれた経路を通ることで固定された経路を進む AGV (Automatic Guided Vehicle) や、カメラなどのセンサーにより障害物を回避しながら自律的に目的地へと移動する AMR (Autonomous Mobile Robot) などがある。これらのロボットの世界市場では 2022 年時点で既に 50 億ドルの売り上げを記録しており、2028 年には 200 億ドル、67 万台の移動ロボットが出荷され、設置ベースは 270 万台まで増加するという予測もある [2]。

これらの工場や物流倉庫内で動くロボットを自律的に制御するためには、障害物に当たらないよう回避しながら運搬先の目的地へと到達する必要がある。しかし、工場や倉庫内には壁や棚などの固定障害物だけでなく、ロボット以外の作業員も作業を行っており、作業場所が変化したり一時的に通路に荷物が置かれたりなどの環境の変化が起こることが考えられる。このような環境変化は事前に分かるものではなく、ロボットが環境変化が起きた場所へ移動し観測を行い初めてわかるものである。

こういった自律的な障害物の回避行動をとるロボットの研究は既に進められている。例えば文献 [3] や [4], [5] では、自律型ロボットが移動する障害物が存在する環境においてそれを回避しながら目的地へと到達可能であることが示されている。特に [4] では経路の導出にニューラルネットワークが使用されており、目の前に障害物が突然現れるといった急激な環境変化にも対応できることが示されている。しかし、障害物を回避することで目的地への移動に要する時間が増加するため、効率的な移動を実現するためには単に障害物を回避するだけでなく障害物がある可能性が低い効率的な経路について考える必要がある。文献 [6] では、これを実現するため自律型ロボット同士が互いに通信しあうことで、お互いを回避するだけでなく無駄な経路を通らないことでその運搬にかかる平均時間を減らすことが可能であることが示されている。しかし、文献 [6] の各ロボットは自身の経路上にある障害物についての情報を共有するだけであり、経路以外についての情報を知ることはできない。そこで、自身の経路以外についての環境情報の収集についても考慮することで、より早く状況を把握しその後の経路設計に活かすことができる可能性があると考えられる。

さらに搬送ロボットにおける荷物の運搬について考えると、障害物だけでなく作業員や荷物の位置についての情報も重要である。従来の搬送ロボットでは、ある箇所の荷物を運び終えた時に次に向

かうべき荷物の場所をロボット自身が把握できていない場合が多く、その場合ロボットはまず運ぶべき荷物の探索を行い、荷物の場所に到着してから目的地までの移動について考え始める必要がある。

これを解決するためには、作業者の作業位置や、次に運搬すべき荷物の場所をはやくに把握しておく必要がある。すなわち、目的地への移動だけでなく作業者の作業位置や、次に運搬すべき荷物の場所などの周囲の状況の把握のための行動が必要となり、状況を把握するための移動という行動と、障害物を回避しながら目的地へと到達するための移動という行動の両立が求められることとなる。

本報告では、荷物の搬送と倉庫内の作業者や搬送すべき荷物の位置などの状況の把握を両立できる自律型搬送ロボットの制御手法を提案する。本手法ではこの制御を能動推論を用いて行う。能動推論では、情報がない状態から観測と行動を繰り返し、自由エネルギーを最小化する状態が移るべき状態とし行動することを人の脳の行動原理とし、これに基づいて状態の推論を行う [7][8]。物流倉庫内で動くロボットは、壁や棚、倉庫内で働く作業者などの障害物の回避と、荷物の運搬のため荷物の置かれている場所と倉庫の出口への往復を同時に実現する必要がある。これは状況が不明であるという情報がない状態から、探索と状況の観測を繰り返し目的地という望ましい場所へと移動を行うという解釈ができ、これは能動推論が考える人の脳の行動原理と一致する。そして、倉庫内の状況把握と目的地への移動という内容は能動推論において統合的に扱うことができる。そこで本報告では、搬送ロボットにおける荷物の搬送先である目的地へと移動する行動と、自身の周囲を観測することにより倉庫内の状況を推測する過程をモデル化することで、搬送ロボットの制御を能動推論により行う手法を提案する。そしてシミュレーション評価により、提案手法を用いてロボットを制御することで倉庫内の状況を早くに把握することができることを示す。

以下で本報告の構成について述べる。まず第2章では能動推論についての説明を行う。そして第3章では提案手法である能動推論に基づく搬送と倉庫内の状況監視を両立する搬送ロボット制御手法について述べ、第4章でその評価結果について述べる。そして最後に第5章でまとめと今後の課題について述べる。

2 能動推論

2.1 能動推論の概要

文献 [1] では、能動推論と呼ばれる理論が提案されている。能動推論では、エージェントがサプライズを最小化するように行動すると仮定している。サプライズとは、エージェントの現在の感覚的観測、すなわち観測が、そのエージェントが好ましい観測、すなわちそのエージェントの健全な状態を維持するものとどの程度異なるかを測定するものである。このエージェントは、観測がどのように生成されるかについての生成モデルを持っており、観測を通して外界についての状態を推論する。また観測を実際に生成するものは生成プロセスと呼ばれ、真の外界についての状態の情報を含む。図 1 は文献 [1] 内で示された生成モデルと生成プロセスの関係である。

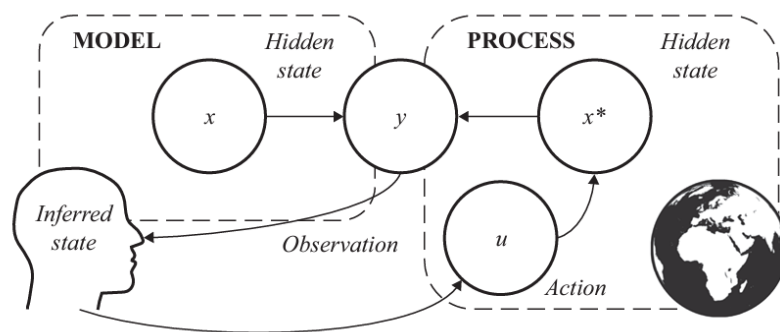


図 1: 生成プロセスと生成モデル, 文献 [1] より引用, p23 図 2.2

エージェントは観測を行い、その観測から生成モデルの状態の推論を行う。そしてサプライズを最小化するように行動し、この行動は生成プロセスの状態を変化させ、新たな観測を生む。

エージェントが持つ典型的な生成モデルは図 2 のような構造である。

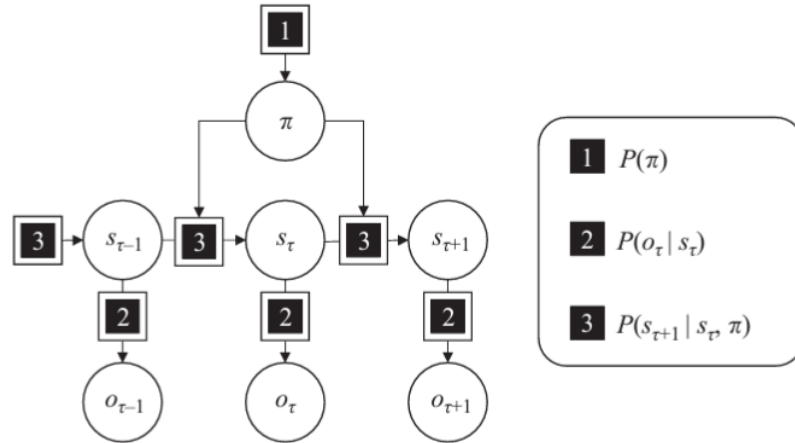


図 2: 生成モデルの構造, 文献 [1] より引用, p70 図 4.3

o は観測を, s は状態を, π はポリシー, すなわち行動の系列を表している. 現在の推論された状態 s_t に基づき観測 o_t が生成され, そこからポリシー π が選択され行動をし, 次の状態 s_{t+1} へと移る.

次に, 状態の推論方法と, ポリシーの選択方法についてである. まず能動推論では, サプライズは変分自由エネルギーと呼ばれる代替指標に置き換わる. 変分自由エネルギーは以下の式で表される.

$$\begin{aligned}
 F[Q, y] &= \underbrace{-E_{Q(x)}[\ln P(y, x)]}_{\text{エネルギー}} - \underbrace{H[Q(x)]}_{\text{エントロピー}} \\
 &= \underbrace{D_{KL}[Q(x) || P(x)]}_{\text{複雑さ}} - \underbrace{E_{Q(x)}[\ln P(y|x)]}_{\text{正確さ}} \\
 &= \underbrace{D_{KL}[Q(x) || P(x|y)]}_{\text{ダイバージェンス}} - \underbrace{\ln P(y)}_{\text{証拠}}
 \end{aligned} \tag{1}$$

ここでは状態を x , 観測を y としている. この変分自由エネルギー F を最小化するように状態の推論を行う. 推論した状態の確率分布は, 推論前が事前信念, 推論後が事後信念と呼ばれる,

また行動の決定においては, 未来のサプライズを最小化するような行動を選択する. これには, 各ポリシーを未来の観測から得るサプライズの観点から評価する期待自由エネルギーと呼ばれる指標

を用いる。期待自由エネルギーは以下の式で表される。

$$\begin{aligned}
G(\pi) &= \underbrace{-E_{Q(\tilde{x}, \tilde{y}|\pi)}[D_{KL}[Q(\tilde{x}|\tilde{y}, \pi)||Q(\tilde{x}, \pi)]]}_{\text{情報利得}} - \underbrace{E_{Q(\tilde{y}|\pi)}[\ln P(\tilde{y}|C)]}_{\text{実利的価値}} \\
&= \underbrace{E_{Q(\tilde{x}|\pi)}[H[P(\tilde{y}|\tilde{x})]]}_{\text{曖昧さの期待値}} + \underbrace{D_{KL}[Q(\tilde{y}|\pi)||P(\tilde{y}|C)]}_{\text{リスク (成果)}} \\
&\leq \underbrace{E_{Q(\tilde{x}|\pi)}[H[P(\tilde{y}|\tilde{x})]]}_{\text{曖昧さの期待値}} + \underbrace{D_{KL}[Q(\tilde{x}|\pi)||P(\tilde{x}|C)]}_{\text{リスク (状態)}} \\
&= \underbrace{-E_{Q(\tilde{x}, \tilde{y}|\pi)}[\ln P(\tilde{y}, \tilde{x}|C)]}_{\text{エネルギーの期待値}} - \underbrace{H[Q(\tilde{x}|\pi)]}_{\text{エントロピー}}
\end{aligned} \tag{2}$$

\tilde{x} , \tilde{y} はそれぞれ状態と観測の系列を表す。この期待自由エネルギー G が小さくなるポリシーの信念が高くなるようにポリシーの事後信念を求める。またポリシーの事後信念から以下の式で行動の決定を行う。

$$\pi(\tau) = \max(\pi_\tau) \tag{3}$$

τ は時間を表しており、その時間での事後信念が最も高いポリシーが選択されることになる。

したがって、能動推論ではまずエージェントは観測を行い、変分自由エネルギー F を計算することで、状態の推論を行う。次に推論した状態の事後信念を基に各ポリシーについての期待自由エネルギー G_π の計算を行い、変分自由エネルギー F とともにポリシーの事後信念を求める。最後にポリシーの事後信念から次の行動を決定し、行動する。この一連の流れを繰り返すことで、エージェントはエージェントが好ましいと考えている状態へと移行していくこととなる。

なお、搬送ロボットにおける目的地への移動と状況把握の統合的な扱いについては期待自由エネルギーで実現されている。期待自由エネルギーは式 3 から以下の式に変形できる。

$$G(\pi) = \underbrace{-H[Q(y_\tau|\pi)] + E_{Q(x_\tau|\pi)}[H[P(y_\tau|x_\tau)]]}_{\text{情報利得}} - \underbrace{\sum_{y_\tau} Q(y_\tau|\pi) \log P(y_\tau|C)}_{\text{選好}} \tag{4}$$

ここで期待自由エネルギーの選好の項で、搬送ロボットにおける障害物の回避や目的地への到達につながる行動を評価することができる。具体的には、選好の確率分布で障害物を観測する場合や目的地から遠い場合の確率を低く、反対に障害物を観測しない場合や目的地から近い場合の確率を高く設定することで、この項は前者の場合に小さく、後者の場合に大きくなる。ポリシー π の事後信念としては、期待自由エネルギーが小さいほどそのポリシーの確率が高くなるので、障害物を観測せず目的地へ近くなる行動を選択しやすくなることになる。また情報利得の項では状況の特定につな

る行動を評価する事ができる。具体的にはまず1つ目の項がその地点での観測の不確実性を表しており、不確実性が大きい時はそこで観測を行えば状況把握につながる事が期待できるため、この項は不確実性が大きい時に小さくなる。また2つ目の項はその地点での状態と観測の関連性を表しており、状態が常に同じ観測を生成する場合その地点に観測の不確実性があれば必ず解消できるためこの項は0に、また乱数生成器のように状態と観測に関連性がない場合は観測に不確実性があっても解消できないためこの項は最大となる。この2つの項で、その地点での観測の不確実性の大きさと、その地点での解消可能な不確実性の大きさの両方を評価することで、状況把握を行うことができるポリシーを選択しやすくなることになる。

2.2 能動推論における生成モデルの設計

一般的に生成モデルの設計は以下のような手順で行う [1].

1. モデル化するシステムの決定
2. 生成モデルの形式の決定
3. 生成モデルの適切な変数の設定

まずモデル化するシステムを決定する。具体的には能動推論を実行するエージェントを決定する。次に生成モデルの形式を決定する。具体的には能動推論を離散時間で実行するか連続時間で実行するかや、モデルの層構造について決定する。モデルが多層であると、異なる時間スケールで変化する変数を含むモデルを作成できる。

最後に生成モデルの変数の設定である。まず設定する変数の導出のため離散時間での変分自由エネルギー、期待自由エネルギーの式変形を行う。

生成モデルは以下のように定義される。

$$P(\tilde{o}, \tilde{s}, \pi) \quad (o: \text{観測}, s: \text{状態}, \pi: \text{ポリシー})$$

また変分自由エネルギーの各項は以下のように変形できる。

$$\begin{aligned}
Q(\tilde{s}|\pi) &= \Pi_\tau Q(s_\tau|\pi) & Q(s_\tau|\pi) &= \text{Cat}(S_{\pi\tau}) \\
P(\tilde{s}|\pi) &= P(s_1)\Pi_\tau P(s_{\tau+1}|s_\tau, \pi) & P(s_{\tau+1}|s_\tau, \pi) &= \text{Cat}(B_{\pi\tau}) & P(s_1) &= \text{Cat}(D) \\
P(\tilde{o}|\tilde{s}, \pi) &= \Pi_\tau P(o_\tau|s_\tau, \pi) & P(o_\tau|s_\tau, \pi) &= \text{Cat}(A) \\
P(\tilde{o}|\pi) &= \Pi_\tau P(o_\tau|\pi) & P(o_\tau|\pi) &= \text{Cat}(O_{\pi\tau})
\end{aligned}$$

Cat はカテゴリカル分布を表す。\$S_{\pi\tau}\$ はポリシーが \$\pi\$、時間が \$\tau\$ の時の状態の確率分布を表すベクトルである。また \$B_{\pi\tau}\$ はポリシーが \$\pi\$、時間が \$\tau\$ の時の状態の遷移確率を表す行列である。列が現在の状態を表し、行が次の状態を表す。また \$D\$ は初期状態を表すベクトルである。これはエージェントの初期位置を表すことになる。また \$A\$ はある状態である観測をする確率である尤度を表す行列である。列が状態を表し、行が観測を表す。

これらの行列を用いてポリシー \$\pi\$ における変分自由エネルギー \$F_\pi\$ を表すと以下ようになる。

$$F_\pi = S_{\pi 1}(\ln S_{\pi 1} - \ln D - \ln A^T O_{\pi 1}) + \sum_{\tau=2} S_{\pi\tau}(\ln S_{\pi\tau} - \ln B_{\pi-1\tau-1} S_{\pi-1\tau-1} - \ln A^T O_{\pi\tau}) \quad (5)$$

また期待自由エネルギーの各項は以下のように変形できる。

$$\begin{aligned}
P(\tilde{o}|\pi) &= \Pi_\tau P(o_\tau|\pi) & P(o_\tau|\pi) &= \text{Cat}(O_{\pi\tau}) \\
P(\tilde{o}|C) &= \text{Cat}(C_\tau) \\
Q(\tilde{s}|\pi) &= \Pi_\tau Q(s_\tau|\pi) & Q(s_\tau|\pi) &= \text{Cat}(S_{\pi\tau}) \\
P(\tilde{o}|\tilde{s}, \pi) &= \Pi_\tau P(o_\tau|s_\tau, \pi) & P(o_\tau|s_\tau, \pi) &= \text{Cat}(A)
\end{aligned}$$

\$O_{\pi\tau}\$ はポリシーが \$\pi\$、時間が \$\tau\$ の時の観測の確率分布を表すベクトルである。また \$C_\tau\$ は時間が \$\tau\$ の時の選好の確率分布を表すベクトルである。この列ベクトルを全時間においてまとめたものを \$C\$ 行列とする。

また以下のような式を定義する。

$$\begin{aligned}
O_{\pi\tau} &= A * S_{\pi\tau} \\
\text{diag}(A^T * \log A) &= H
\end{aligned}$$

これらの行列を用いてポリシー \$\pi\$ における期待自由エネルギー \$G_\pi\$ を表すと以下ようになる。

$$G_\pi = H S_{\pi\tau} + A * S_{\pi\tau}(\log A * S_{\pi\tau} - \log C_\tau) \quad (6)$$

ここまでで変分自由エネルギー，期待自由エネルギーを4つの変数行列 A, B, C, D を用いて表すことができた。

さらに，式変形した変分自由エネルギーをもとに状態の事後信念を以下の式で求めることができる。

$$\begin{aligned}
 s_{\pi\tau} &= \sigma(v_{\pi\tau}) \\
 \dot{v}_{\pi\tau} &= \epsilon_{\pi\tau} \\
 \epsilon_{\pi\tau} &= \ln A * o_{\tau} + \frac{1}{2}(\ln(B_{\pi\tau}s_{\pi\tau-1}) + \ln(B_{\pi\tau+1}^{\dagger}s_{\pi\tau+1})) - \ln s_{\pi\tau} \\
 B_{\pi\tau}^{\dagger} &\propto B_{\pi\tau}^T
 \end{aligned} \tag{7}$$

また，ポリシーの事前信念を指定する E 行列を新たに設定することで，式変形した期待自由エネルギーと変分自由エネルギーをもとにポリシーの事後信念を以下の式で求めることができる。

$$\pi = \sigma(\ln E - F - G) \tag{8}$$

したがって，生成モデル $P(\tilde{o}, \tilde{s}, \pi)$ において5つの変数行列 A, B, C, D, E を設定することで，状態やポリシーの事後信念を計算し行動を決定することができるため，この生成モデルを持つエージェントが能動推論を実行することができることになる。

2.3 能動推論を応用した従来手法

文献 [9] では生物が行う周囲の観測や目的地までの経路計画などが能動推論による定式化で表すことができることから，これをロボットに対し適用した能動推論を用いて行動計画を行うロボットについて提案されている。しかし，提案されている手法の行動計画では従来の搬送ロボットと同様に障害物の回避と目的地への到達を重視した内容となっており，状況把握については考慮されていない。そこで本報告では状況把握について考慮した手法を提案する。

3 能動推論に基づく搬送と倉庫内の状況監視を両立する

搬送ロボット制御手法

3.1 提案の概要

物流倉庫内で動く搬送ロボットの自動制御において、能動推論に基づいた行動決定を行うことにより荷物の搬送と倉庫内の作業員や搬送すべき荷物の位置などの状況の把握を両立させる。具体的には、能動推論での対象のエージェントをロボットに、観測を作業員や障害物の有無、状態をロボットの位置や倉庫の状況、ポリシーを移動として生成モデルを定義し、生成モデルで尤度、状態の遷移確率、選好、初期状態、ポリシーの事前信念を表す5つの変数行列を対象の倉庫内の作業員などの障害物や目的地の位置から設定し、能動推論を用いて行動を決定する。また、ロボットが目的地に着くと再度生成モデルの変数行列を設定しなおす。

設計する生成モデルは図で表すと図3のようになる。

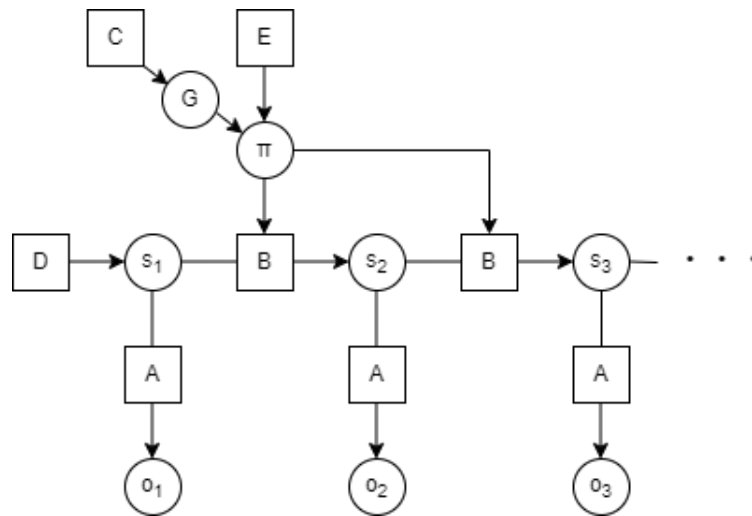


図 3: 設計した生成モデル

まず初期状態を表す D ベクトルから初期状態である s_1 を求め、尤度を表す A 行列から観測を行い o_1 を求める。そして変分自由エネルギーを計算し第 2.2 節の式 7 を用いて状態の事後信念を求める。この際ロボットは倉庫内の状況についての情報を状態として持っており、観測から変分自由エネルギーを計算し、第 2.2 節の式 7 を用いてこの倉庫内の状況についての状態の事後信念を求めることで倉庫内の状況の推定を行う。次に選好を表す C 行列から期待自由エネルギー G を求め、ポリシー

の事前信念を表す E 行列とともに第 2.2 節の式 8 を用いてポリシーの事後信念 π を求める。そして第 2.1 節の式 3 を用いてポリシーを選択し、対応する行動をした結果の遷移確率を表す B 行列から次の状態 s_2 を求め、これを繰り返す。

3.2 提案手法における生成モデルの詳細

本節では提案手法における生成モデルの詳細について述べる。

まずモデル化するシステムを決定する。具体的には工場や物流倉庫内で動くロボットを能動推論を実行するエージェントとみなす。

次に生成モデルの形式を決定する。具体的には能動推論を離散時間で実行するか連続時間で実行するかや、モデルの層構造について決定する。モデルが多層であると、異なる時間スケールで変化する変数を含むモデルを作成できる。ここでは離散時間で実行する単層のモデルとする。

最後に生成モデルの変数の設定である。ここからは各項に分けて説明する。

モダリティ数と観測内容

観測するモダリティは 2 つである。このモダリティの数だけ尤度を表す A 行列が存在することになる。1 つ目のモダリティは自身の位置である。このモデルでは倉庫をグリッド上に分割して考えるものとし、自分がグリッド上に分割されたどの位置にいるかをセンサー等で判別できるものとし、それを観測とする。観測の種類としては倉庫内のグリッド上に分割されたマス数だけあることになる。2 つ目のモダリティは周囲の障害物の有無である。グリッド上に分割されたマスの自身の位置から上下左右 3 マスに対して、カメラ等のセンサーを用いて障害物の有無を判別できるものとし、それを観測とする。観測の種類としては前後左右の 4 方向それぞれの障害物の有無に対し 2^4 の 16 種類と障害物上にいるという観測を合わせた 17 種類である。

状態の要因の数と状態の内容

状態の要因の数は 2 つである。1 つ目の要因は自身の位置である。観測をもとに自分が倉庫内のグリッド上に分割されたどの位置にいるかを確率としてもつ。とりうる状態数は倉庫内のグリッド上に分割されたマス数だけあることになる。2 つ目の要因は作業者の位置という倉庫内の状況である。

ロボットは、倉庫内の通路にいる作業者の位置の組み合わせの全通りの候補の確率を状態として持っており、観測をもとに可能性のある位置の組み合わせを求め倉庫の作業者の位置に関する状況の信念として更新する。

ポリシーの内容

ポリシーはロボットの1ステップ当たりの行動の系列を表している。1ステップ当たりの行動としては1マス分前後左右に進むことと、その場で静止する5種類がある。考えられるポリシー数としては5の全体の移動にかかったステップ数乗となり、その数が膨大であるため本手法では考えるポリシーを限定する。限定方法としては、現在ロボットが向かおうとしている地点に対し、そこに行くために大きく遠回りをする必要がないという仮定を置き、ロボットの現在地を除いた全ての地点に対し、同じ地点を含むことを許した最短経路から短い順の決められた個数の経路に対応するポリシーのみを考えることとする。同じ地点を含むことを許した最短経路から短い順の決められた個数の経路は、Eppstein's algorithm と呼ばれるグラフアルゴリズムを用いる [10]。この Eppstein's algorithm は多重有向グラフにおいて閉路を含めた最短経路から短い順の決められた個数の経路を導出するアルゴリズムであり、倉庫を壁や柵以外の通路のみが連結したグラフとし Eppstein's algorithm を用いることで、同じ地点を含むことを許した最短経路から短い順の決められた個数の経路を導出する。これにより、考えるポリシー数を経路の決められた個数と通路の全マスの積の値に制限できる。なお、この考えるポリシーは毎ステップ変化するため毎ステップ指定しなおす必要がある。

尤度

尤度を表す A 行列について述べる。まず自分の位置に関する A 行列については、ロボットが自分の位置に関する情報をセンサーなどを用いて観測できるものとする、各位置、すなわち状態に対して観測が一对一で対応することになる。つまりこの A 行列は位置に関する状態の数の大きさの単位行列となる。次に障害物の有無に関する A 行列については、ロボットが周囲の障害物の有無に関する情報をセンサーなどを用いて観測できるものとする、各位置、すなわち状態に対して障害物を観測する確率を指定することになる。観測の種類は大きく分けて障害物が存在しない、障害物が存在する、または障害物上であるという3種類に分けられ、障害物が壁や柵などの移動しない障害物の場

合は確率が1であり、人などの移動する障害物の場合は等確率として0.5となる。よってある方向に対し自身の位置から3マス以内に移動しない障害物がある場合は1、作業者がいる可能性があるマスが1つ含まれている場合は0.5、2つ含まれている場合は $1 - 0.5 * 0.5$ の0.75、障害物が含まれない場合は0となる。また障害物上にあるという観測に対しては障害物上の位置に対して1を指定する。またこのA行列の大きさは障害物を認識する角度の数によって変わる。センサーが360度観測できる場合角度をいくつに分割するかによって行数は変化する。例えば前後左右の4つに分割する場合は行数は 2^4 の16となる。

遷移確率

遷移確率を表すB行列について述べる。自分の位置に関するB行列については、まずB行列は各ポリシーに対し1つ存在し、対応するポリシーを行ったときに遷移先となる状態の確率を指定する。ロボットの移動がポリシー通りに必ず成功するとすると、各位置の状態から対応するポリシーを行った時の移動先の位置の状態への遷移確率が1、それ以外の状態への遷移確率が0となるように指定する。ここでは各ポリシーではなく1ステップ当たりの行動に対して1つB行列を設定する。次に倉庫の作業者の位置に関するB行列については、ロボットは作業者の位置に関する状態は変化しないという仮定を置き作業者の位置が変化しなかった場合の経路を考えて進み、変化している場合観測によってその都度信念を更新することになるため作業者の位置に関する遷移前の状態と遷移先の状態は同じである。したがってこのB行列は考える作業者の位置の組み合わせの大きさの単位行列となる。

選好

選好を表すC行列について述べる。このC行列はモダリティに関連するため尤度を表すA行列と対応しており、各モダリティの各観測についての選好を表す。まず自分の位置に関するC行列については、ロボットを目的地へと誘導するため、目的地であるという観測が、目的地から遠い場所である観測よりも起きやすいという確率分布とする。具体的には行列の各要素を以下の式で定義する。

$$C_{s\tau} = pr \left(1 - \frac{D(s)}{\max(D(s))} \right)$$

ここで $D(s)$ は目的地と状態 s が表す位置とのユークリッド距離である。また pr は目的地への選好の強さを表すパラメータである。行列の値としては目的地に対応する要素が pr 、目的地から最も遠

い場所に対応する要素が 0, それ以外の要素は距離によって比例するものとなる。また, 目的地に対応する要素にのみ以下のようにパラメータ α を加える。

$$C_{s=\text{目的地}r} = pr + \alpha$$

このパラメータ α は評価時にいくつかの数値を設定し評価する。そして各列ベクトルをソフトマックス関数を用いて正規化し確率分布へと変換する。またこの選好は時間によって変化しないものとする。すなわち全ての列ベクトルは等しいものとする。次に障害物の有無に関する C 行列を設定する。障害物の回避のため, 障害物上でないという観測が, 障害物上であるという観測よりも起きやすいという確率分布とする。具体的には障害物上でないという観測に対応する要素を 0, 障害物上であるという観測に対応する要素を障害物への嫌悪の強さを表すパラメータ av とし, 各列ベクトルをソフトマックス関数を用いて正規化し確率分布へと変換する。またこの選好も自分の位置に関する C 行列と同様に時間によって変化しないものとする。すなわち全ての列ベクトルは等しいものとする。

初期状態

初期状態を表す D ベクトルについて述べる。この D ベクトルは状態に関連するため遷移確率を表す B 行列と対応しており, 各要因の各状態が初期状態である確率を表す。まず自分の位置に関する D ベクトルについては, 初期状態であるためロボットの初期位置に対応する状態に確率 1 を, それ以外の状態に 0 を指定する。次に作業者の位置に関する D ベクトルについては, 初めの移動では情報がないため各状態を等確率とする。2 回目以降の移動では, 前回の移動の目的地に着いた最後のステップでの倉庫の状況に関する状態の信念をそのまま指定する。このようにすることで各移動間で作業者の位置に関する状態の信念を引き継ぐこととなる。

ポリシーの事前信念

ポリシーの事前信念を表す E 行列について述べる。この E 行列は初期段階でエージェントがどのポリシーをとりやすいかを表す。初期段階では情報がないためポリシーの事前確率は各ポリシーを等確率とする。

4 提案手法の評価

4.1 設定

4.1.1 環境

評価時の倉庫の内部構造は図4の通りである。

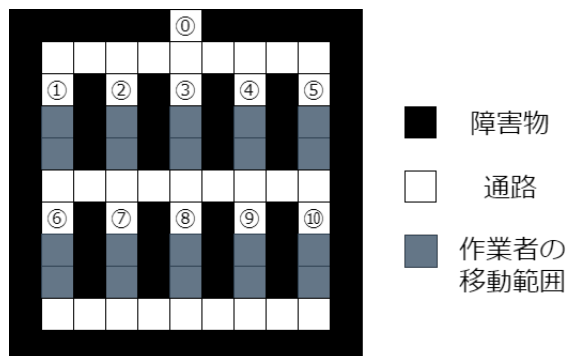


図4: 倉庫の内部構造

倉庫内をグリッド上に分割して考える。1ステップで作業者やロボットはグリッドの1マス分進むことができる。なお黒色の四角が壁や荷物がある棚を表しており、白色の四角がロボットや作業者が通ることのできる通路を表している。また灰色の四角は作業者の移動範囲を表しており、作業者は10本ある上下方向の通路のいずれかにいるものとする。また1番から10番は作業者の荷物の積み下ろし場所、0番は倉庫の出口を表している。ロボットは1番から10番のどこかで1人の作業者に対して荷物の運搬を行う。

また、作業者は移動する場合と移動しない場合の両方について考える。まず移動しない場合についての作業者や荷物の位置は図5の通りである。

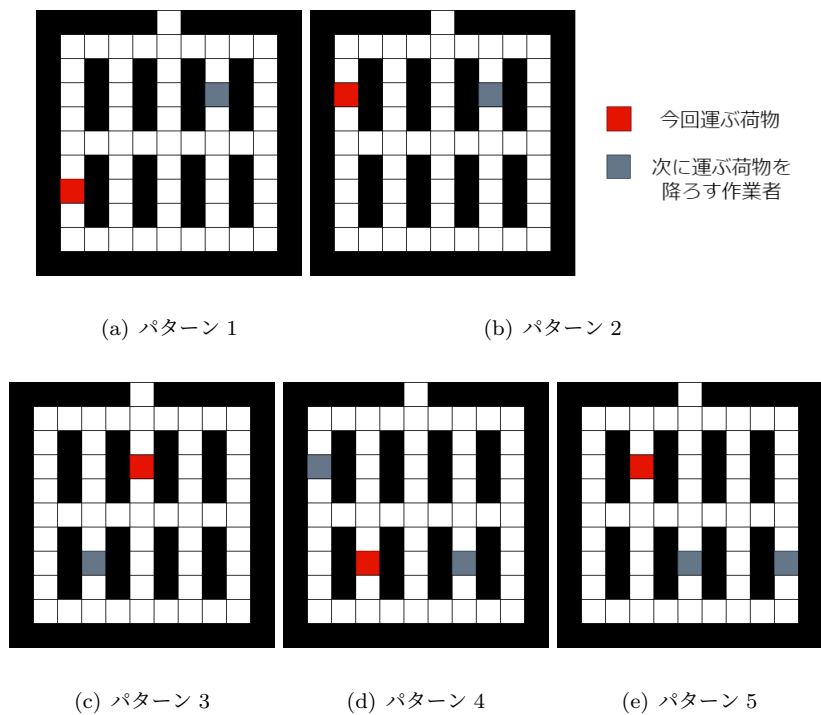


図 5: 作業員が移動しない場合の作業員や荷物の位置の設定

ロボットは初め作業員がおり荷物がある赤色の四角の地点の1つ上のマスにおり、倉庫の出口まで荷物を3回にわたって運ぶ。また灰色の四角は次に運ぶべき荷物の位置である。また図5(a), (b), (c)では倉庫内の作業員は2人であり、作業員の位置の考えられる状況の組み合わせは45通りとなり作業員の位置に関する状態の数も45個となる。また図5(d), (e)では倉庫内の作業員は3人であり、2人の場合と同様に同様に作業員の位置に関する状態の数は120個となる。

また作業員が移動する場合の作業員や荷物の位置は図6の通りである。

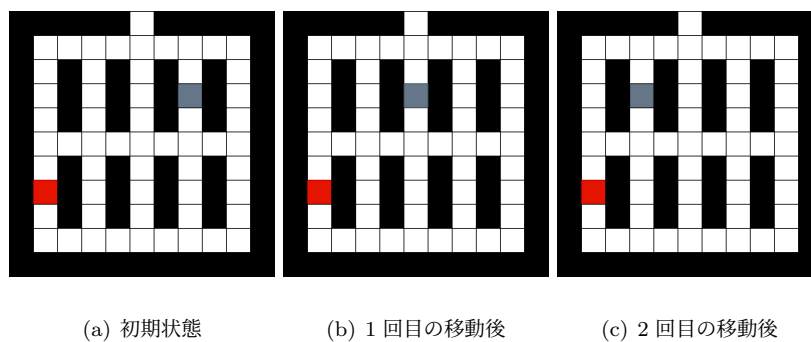


図 6: 作業員が移動する場合の作業員や荷物の位置の設定

図 6(a) の位置から図 6(b), 図 6(c) のように次に運ぶ荷物の積み下ろしを行う作業者が動き位置が変化する。作業者の移動タイミングとしてはロボットが 1 回目に荷物を倉庫の出口へと運び荷物へと戻る移動の初めの 1 回, 荷物と倉庫の出口を 1 往復し 2 つ目の荷物を倉庫の出口へと運ぶ途中の 1 回, 1 回目に荷物を倉庫の出口へと運び荷物へと戻る移動の途中と 2 回目に荷物を倉庫の出口へと運び荷物へと戻る移動の初めの 2 回の 3 つのパターンとする。

この 5 つの作業者が動かない場合と 3 つの作業者が動く場合の計 8 つの状況に対しシミュレーションを行う。

4.1.2 比較対象

提案手法の評価としては、行動決定に能動推論を用いるロボットと、能動推論を用いないロボットの 2 つのロボットに対し同じ倉庫の内部構造、作業者の動きをする場所でシミュレーションを行い、進んだ経路や作業者の位置に関する状態の信念の変化を比較することで評価を行うこととする。ここで行動決定に能動推論を用いないロボットについて説明する。まず行動決定の方法は、目的地までの最短経路を行動決定に能動推論を用いるロボットと同様に倉庫をグラフ化し Eppstein's algorithm を用いて求め、求めた経路に対応する行動の系列の初めの行動を次にとる行動とする。この倉庫のグラフ化の際、壁や柵などの移動しない障害物がある位置についてはグラフのノードに含まないようにすることで移動しない障害物を回避できるようにする。次に移動する障害物である作業者の回避方法については、能動推論を用いて行動決定を行うロボットと同様に観測を行い、作業者を観測すると作業者がいる位置に対応するグラフのノードを削除しその位置を通る経路を考える経路から除くことで回避できるようにする。また評価のために能動推論を用いて行動決定を行うロボットでの作業者の位置に関する状態の信念に対応する確率分布の計算も行う。計算方法としては、まず求めるものをその時点の観測をもとにした作業者の位置の組み合わせの実際の倉庫の状況と一致している確率とする。これはその時点での観測と一致している作業者の位置の組み合わせを、作業者の位置の組み合わせの総数で割ることで求めることができる。なお以後本報告ではこの能動推論を用いずに行動決定を行うロボットにおけるその時点の観測をもとにした作業者の位置の組み合わせの実際の倉庫の状況と一致している確率を、表記を揃えるため能動推論での作業者の位置に関する状態の信念と同義として扱う。また以前の観測と一致しなかったり、作業者が実際の人数よりも多い、または少ないという観測になった場合作業者が移動したとして現在までの観測や行動決定の際に使

用するグラフを初期化し、新たに観測を始める。

したがって評価の際にあらわれる能動推論を用いるロボットとの違いは、状況把握のための移動を行わないことと、作業者の位置に関する状態の信念の計算方法である。

4.1.3 評価指標

第 4.1.1 項で述べた 8 つの状況に対し、能動推論を用いるロボットの目的地への選好のパラメータを変化させながら能動推論を用いるロボットと用いないロボットでシミュレーションを行い、進んだ経路や作業者の位置に関する状態の信念の変化、また状況把握による経路長への影響を比較し評価を行う。作業者の位置に関する状態の信念は第 3.2 節で設計した生成モデルの 2 つ目の状態の確率分布であり、実際の作業者の位置に対応する状態の信念が高くなるほど、状況把握を行うことができているといえることになる。

能動推論を用いるロボットのパラメータについては、第 3.2 節の選好の変数行列で述べた 3 つのパラメータ pr , α , av を設定する。まず pr については、その値を $\frac{1}{5}$ と固定する。すなわち目的地であるという観測が、目的地から遠い場所である観測よりも $e^{\frac{1}{5}}$ 倍起きやすいということになる。この $\frac{1}{5}$ という数値は、倉庫内の状況把握に対応する情報利得から計算された期待自由エネルギーの値に対し、目的地との距離によるこの C 行列の選好から計算された期待自由エネルギーの値が情報利得が無視されるほど小さくならず、また情報利得が重視されすぎないほど大きくもならない値をいくつか検証した結果選ばれた値である。次に av については、その値を -1000 と固定する。すなわち障害物上でないという観測が、障害物上であるという観測よりも e^{1000} 倍起きやすいということになる。この -1000 という数値は、障害物上である観測をもとに計算された期待自由エネルギーの値がその観測を含むポリシーが選択されなくなるほど充分大きくなるように設定された値である。最後に α については、その値を変化させる。具体的には $\alpha=0, 0.5, 1, 2, 3, 5$ の 6 種類について評価を行う。なお以後本報告ではこの α を選好パラメータ α と呼ぶ。

4.2 評価結果

図 7 から図 11 は、それぞれ図 5 の場合で通った経路に違いが出た場合の経路を矢印で示している。

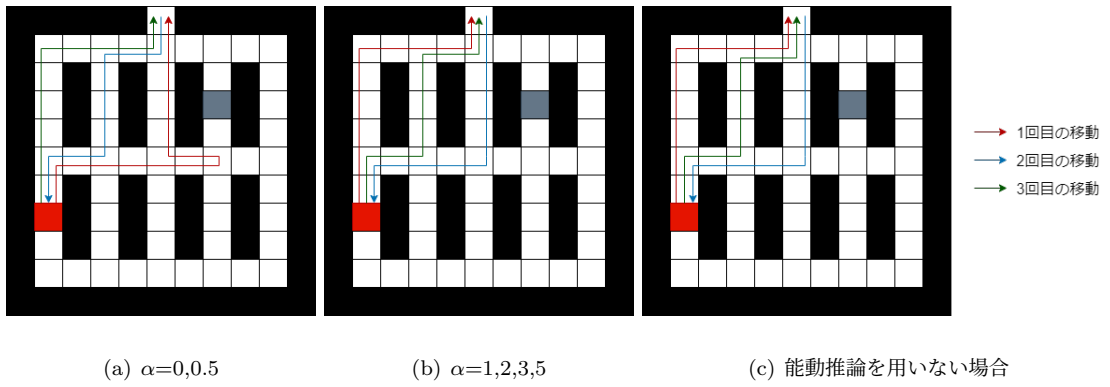


図 7: 図 5(a) でロボットが進んだ経路

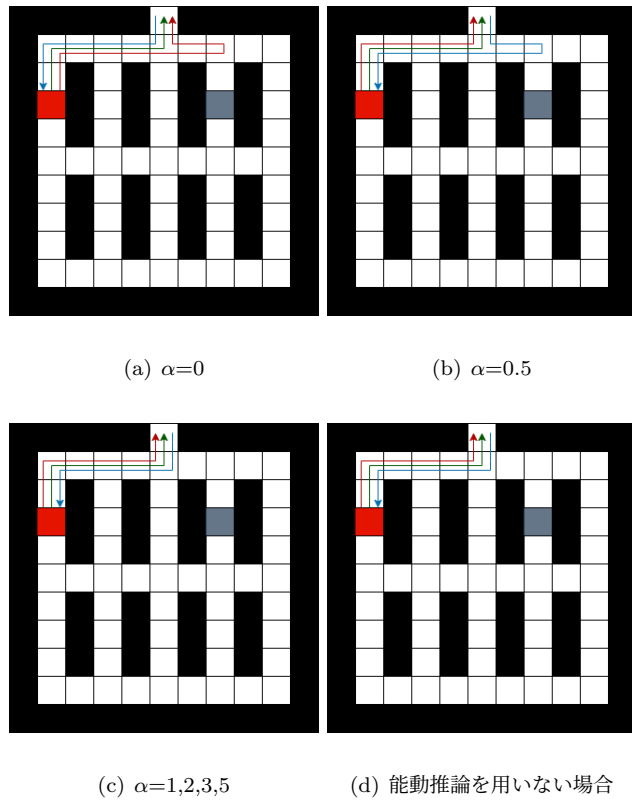
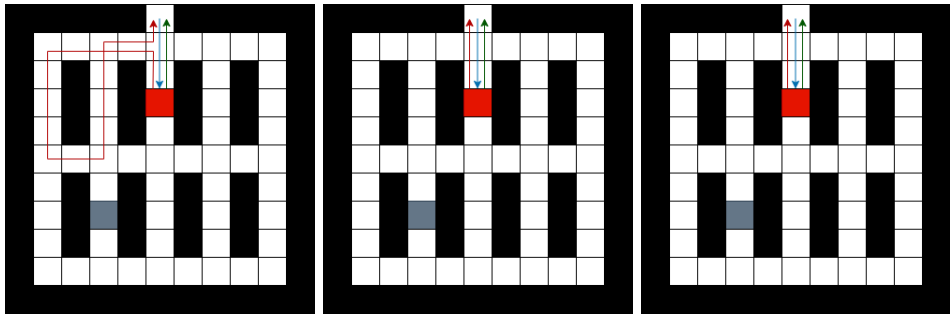


図 8: 図 5(b) でロボットが進んだ経路

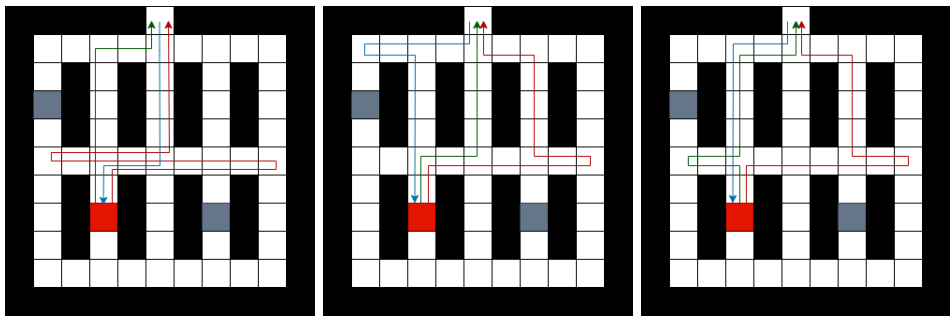


(a) $\alpha=0$

(b) $\alpha=0.5, 1, 2, 3, 5$

(c) 能動推論を用いない場合

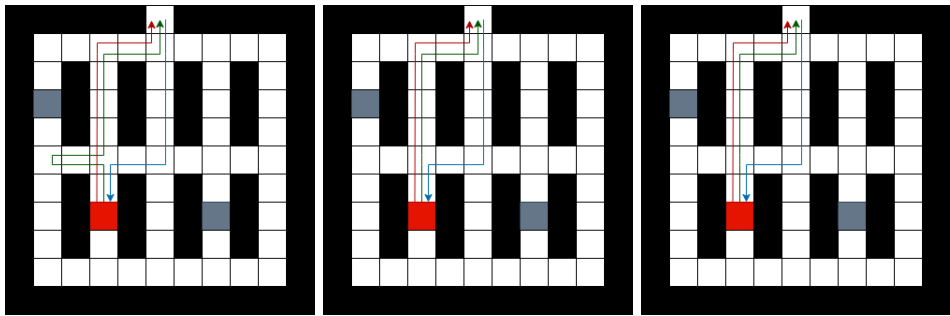
図 9: 図 5(c) でロボットが進んだ経路



(a) $\alpha=0, 0.5$

(b) $\alpha=1$

(c) $\alpha=2$

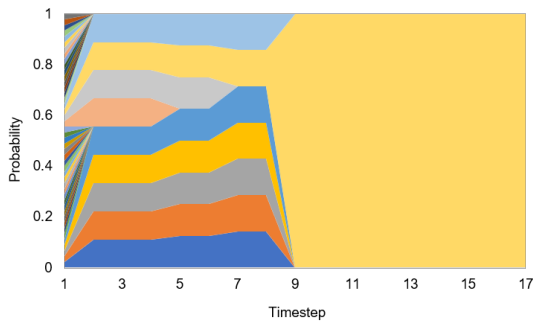


(d) $\alpha=3$

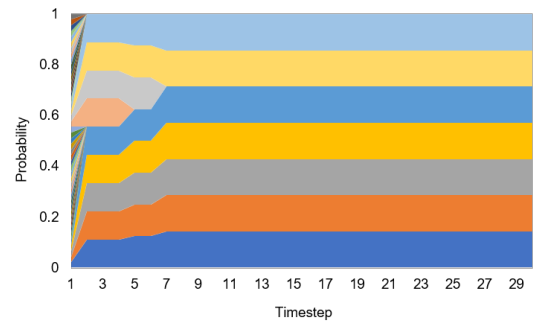
(e) $\alpha=5$

(f) 能動推論を用いない場合

図 10: 図 5(d) でロボットが進んだ経路

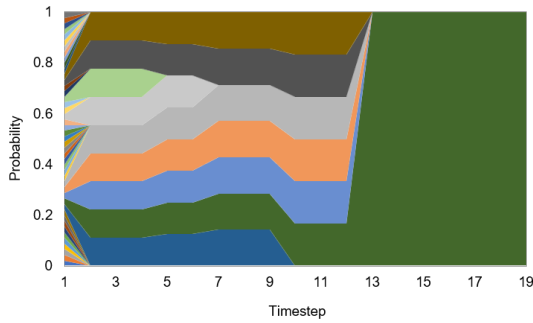


(a) 能動推論を用いた場合 ($\alpha=0$)

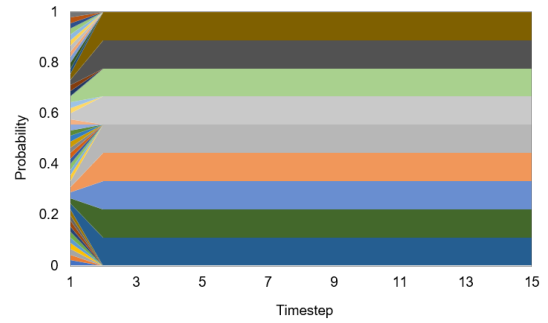


(b) 能動推論を用いない場合

図 13: 作業者が移動しない場合での作業者の位置に関する状態の信念の変化 (図 5(b))

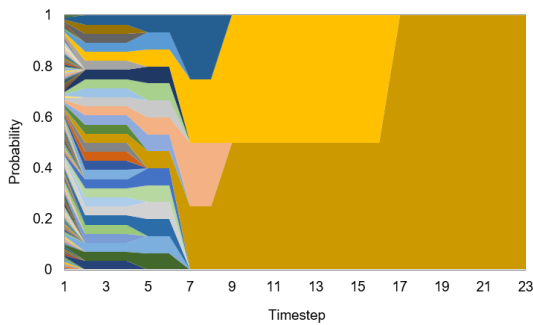


(a) 能動推論を用いた場合 ($\alpha=0$)

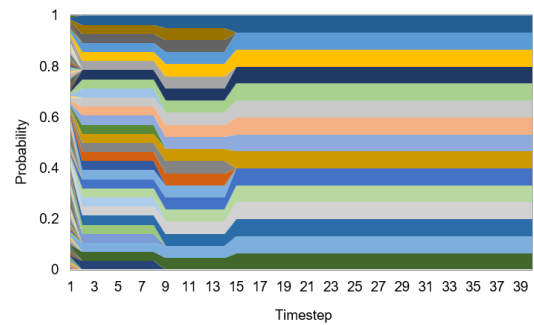


(b) 能動推論を用いない場合

図 14: 作業者が移動しない場合での作業者の位置に関する状態の信念の変化 (図 5(c))

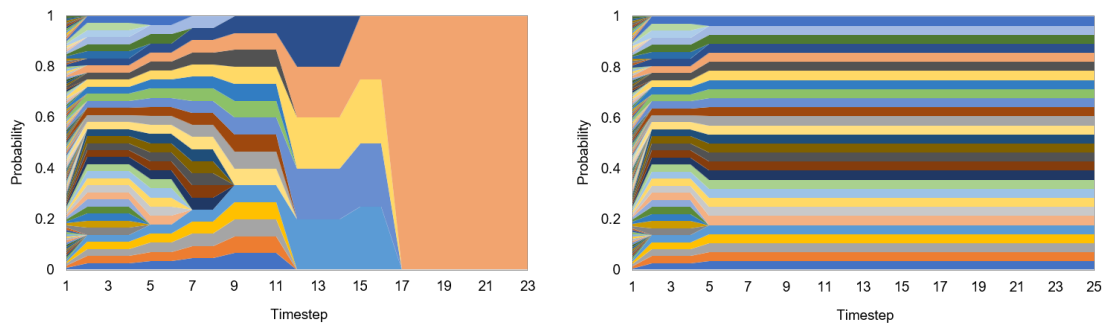


(a) 能動推論を用いた場合 ($\alpha=0$)



(b) 能動推論を用いない場合

図 15: 作業者が移動しない場合での作業者の位置に関する状態の信念の変化 (図 5(d))



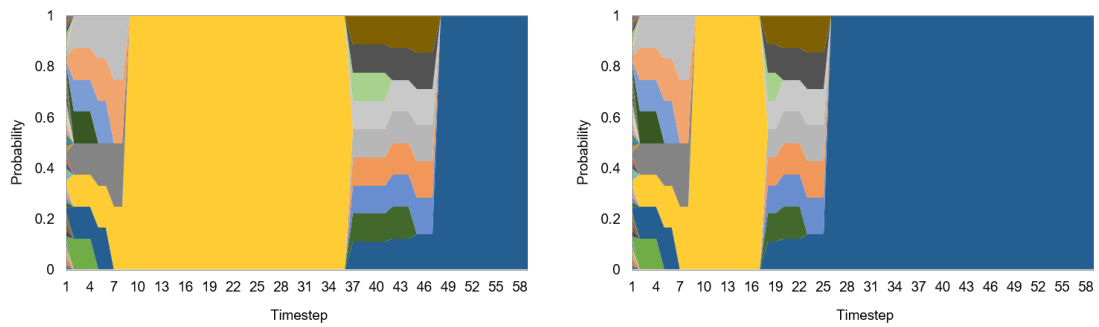
(a) 能動推論を用いた場合 ($\alpha=0$)

(b) 能動推論を用いない場合

図 16: 作業者が移動しない場合での作業者の位置に関する状態の信念の変化 (図 5(e))

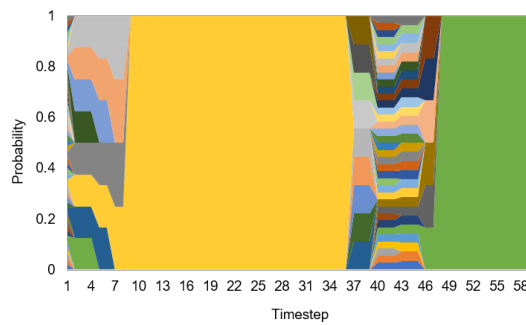
各グラフでは、各色で区切られた領域が図 12, 13, 14 では 45 通り、図 15, 16 では 120 通りある作業者の位置の組み合わせのいずれかの状況を表す状態の信念と対応しており、面の縦軸における幅が確率を表している。また図 12 から図 16 のいずれの場合でも (a) では 1 になっている信念が実際の倉庫の作業者の位置に対応した状態の信念であり、(b) ではどの状態の信念も 1 まで大きくならないことから、能動推論を行動決定に用いて状況把握を行うことで、作業者の位置に関する状態の信念を更新し作業者の位置を把握できていることがわかる。

また図 17 は図 6 の作業者が動く場合での能動推論を用いたロボットの作業者の位置に関する状態の信念の変化を表したグラフである。図 17(a) はロボットが 1 回目に荷物を倉庫の出口へと運び荷物へと戻る移動の初めに移動する場合、図 17(b) は荷物と倉庫の出口を 1 往復し 2 つ目の荷物を倉庫の出口へと運ぶ途中に移動する場合、図 17(c) は 1 回目に荷物を倉庫の出口へと運び荷物へと戻る移動の途中と 2 回目に荷物を倉庫の出口へと運び荷物へと戻る移動の初めに移動する場合である。



(a) パターン 1

(b) パターン 2



(c) パターン 3

図 17: 作業者が移動する場合での作業者の位置に関する状態の信念の変化 (図 6)

図 17(a) では 31 ステップ目で、(b) では 16 ステップ目で、(c) では 22 ステップ目と 38 ステップ目で作業者は移動している。図 6(a) から 6(c) のいずれの場合でも作業者の位置が変化していることを観測すると、それに伴って信念を更新できていることが分かる。

また表 1 は図 5 の (a), (b), (c), (d), (e) の場合でロボットが荷物を 3 回倉庫の出口へと運んだ際に進んだ合計の経路長を表している。能動推論を用いて行動決定を行うロボットでは、状況把握のための移動を行うことで能動推論を用いずに行動決定を行うロボットが進んだ経路長よりも経路長が長くなっていることが分かる。

表 1: ロボットが移動した経路長の合計

	能動推論を用いるロボット						能動推論を用いないロボット
	$\alpha=0$	$\alpha=0.5$	$\alpha=1$	$\alpha=2$	$\alpha=3$	$\alpha=5$	
図 5(a)	59	59	55	55	55	55	55
図 5(b)	39	39	35	35	35	35	35
図 5(c)	31	15	15	15	15	15	15
図 5(d)	61	61	57	57	49	45	45
図 5(e)	41	41	25	25	25	25	25

4.3 考察

第 4.3 節で示した図 7 から図 11 より、能動推論を用いて行動決定を行うロボットでは、選好パラメータ α が小さい時は積極的に状況把握のための移動を行い、 α が大きくなるにつれて自身の位置と目的地の位置が近い場合では状況把握のための移動を行わなくなり、最終的に $\alpha=5$ の場合では全ての場合で状況把握のための移動を行わず能動推論を用いずに行動決定を行うロボットと同じ経路を通ることが分かる。

また表 2 は、能動推論を用いずに行動決定を行うロボットが進んだ経路長に対する、各選好パラメータ α での能動推論を用いて行動決定を行うロボットの経路長の増加率を表している。また表 3 は能動推論を用いて行動決定を行うロボットの状況把握、すなわち作業者の位置に関するある状態の信念が 1 になるまでにかかる時間を表している。最後まで 1 にならない場合は n/a としている。

表 2: 能動推論を用いて行動決定を行うロボットの経路長の増加率

	$\alpha=0$	$\alpha=0.5$	$\alpha=1$	$\alpha=2$	$\alpha=3$	$\alpha=5$
図 5(a)	+7.2%	+7.2%	+0%	+0%	+0%	+0%
図 5(b)	+11%	+11%	+0%	+0%	+0%	+0%
図 5(c)	+107%	+0%	+0%	+0%	+0%	+0%
図 5(d)	+36%	+36%	+27%	+27%	+8.9%	+0%
図 5(e)	+64%	+64%	+0%	+0%	+0%	+0%

表 3: 能動推論を用いて行動決定を行うロボットの状況把握にかかる時間（ステップ数）

	$\alpha=0$	$\alpha=0.5$	$\alpha=1$	$\alpha=2$	$\alpha=3$	$\alpha=5$
図 5(a)	8	8	n/a	n/a	n/a	n/a
図 5(b)	8	10	n/a	n/a	n/a	n/a
図 5(c)	12	n/a	n/a	n/a	n/a	n/a
図 5(d)	16	16	22	28	n/a	n/a
図 5(e)	16	16	n/a	n/a	n/a	n/a

まず表 2 では図 5(c) や (e) などの運ぶ荷物の位置が目的地に近い場合進んだ経路長の合計が短く、また状況把握を行うためにかかる経路長も長くなるため増加率が大きくなっている。また α が大きくなるにつれて増加率が下がる、すなわち状況把握を行わなくなることが分かる。また表 3 では状況把握を行って作業者の位置に関するある状態の信念が 1 になる、すなわち完全に作業者の位置が把握できる場合、 α が大きくなるにつれて把握できるまでにかかる時間が長くなることがわかる。これらの表 2, 3 から、 α の値が小さいと状況把握を行いやすくなりかつ状況把握にかかる時間も短い反面、経路長が状況把握を行わない場合に比べ大きく増加する可能性があり、反対に α の値が大きいと状況把握を行いにくくなり状況把握にかかる時間が長い反面、経路長は状況把握を行わない場合に比べ大きく増加することはないという状況把握と経路長がトレードオフの関係にあることが分かる。したがってこの例では必ず状況把握をしたい場合は $\alpha=0$ 、状況把握もしつつ効率に大きな影響を与えない場合は $\alpha=0.5$ から 1 にするなど、状況把握を重視する度合いによってこの選好パラメータ α を変更することで、提案手法の目的である荷物の搬送と状況把握の両立を実現できることがわかる。

今回設計した提案手法では、物流ロボットの行動決定に能動推論を用いることで荷物の搬送と倉庫内の状況監視を両立することを目的に設計したため、その計算速度については考慮していない。実際に、CPU が Core i7-8550U、メモリが 16G である計算機環境において、図 5(a) の場合の考えられる状況の組み合わせが 45 通りである場合に 1 ステップ当たりの行動決定に要する計算時間が約 9 秒であるのに比べ、図 5(d) の場合の考えられる状況の組み合わせが 120 通りである場合では約 18 秒と大きく増加した。実際の倉庫は評価時の倉庫の内部構造よりも大きく、考えられる状況の組み合わせもより多くなると考えられる。したがって今後は、実用に向けた計算速度の向上について取り組む。また、倉庫内の状況把握による具体的な効果については未検証のため、実際に得られる効果について検証する。そして、実世界をロボットがカメラなどのセンサーで認識し、仮想空間上で生成モデルを

対応させることで、デジタルツインと呼ばれる技術との統合について取り組む。

5 おわりに

本報告では、物流倉庫内で動く搬送ロボットの自動制御において、能動推論に基づいた行動決定を行うことにより荷物の搬送と倉庫内の作業員や搬送すべき荷物の位置などの状況の把握を両立させる制御手法を提案した。能動推論における生成モデルを設計しロボットの行動決定に適用することで倉庫内の状況把握を行い、それに伴い作業員の位置に関する状態の信念を更新できることを示した。また、状況把握を重視する度合いによってパラメータを変化させることで、荷物の搬送と状況把握の両立を実現できることを示した。今回設計した提案手法では、物流ロボットの行動決定に能動推論を用いることで荷物の搬送と倉庫内の状況監視を両立することを目的に設計したため、その計算速度については考慮していない。今後は、実用に向けた計算速度の向上について取り組む。また、倉庫内の状況把握による具体的な効果については未検証のため、実際に得られる効果について検証する。そして、実世界をロボットがカメラなどのセンサーで認識し、仮想空間上で生成モデルを対応させることで、デジタルツインと呼ばれる技術との統合について取り組む。

謝辞

本報告を終えるにあたり，大阪大学大学院情報科学研究科の村田正幸教授には日頃より多大なご指導を賜りましたこと深謝いたします。また，大阪大学サイバーメディアセンターの大下裕一准教授には本研究や本論文の執筆に至るまで日頃よりご指導やご助言を下さいましたこと心より厚く感謝いたします。最後に，大阪大学大学院情報科学研究科村田研究室の皆様には本研究の遂行および本論文作成にあたり多大なご助言，ご協力頂きました。ここに感謝の意を表します。

参考文献

- [1] Thomas Parr, Giovanni Pezzulo, Karl J. Friston, "*Active Inference - The Free Energy Principle in Mind, Brain, and Behavior*". The MIT Press, 2022.
- [2] LogisticsIQ, "AGV (Automated Guided Vehicles) and AMR (Autonomous Mobile Robots) Market Opportunity worth ~\$20B by 2028 with an installed base of 2.7 Million Robots - Driven by Logistics & Manufacturing (4th Edition)," 2023.
- [3] Pfeiffer, M., Schaeuble, M., Nieto, J., Siegwart, R. Cadena, C., "From perception to decision: A data-driven approach to end-to-end motion planning for autonomous ground robots," in *Proceedings of 2017 IEEE International Conference on Robotics and Automation*, May 2017.
- [4] Alan Kunz Cechinel, Anderson Luiz Fernandes Perez, Patricia D. M. Plentz, Edson R. De Pieri, "Autonomous mobile robot using distance and priority as logistics task cost," in *Proceedings of IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society*, Oct. 2020.
- [5] André Lourenço, Francisco Marques, Ricardo Mendonça, Eduardo Pinto, José Barata, "On the design of the ROBO-PARTNER Intra-factory logistics autonomous robot," in *Proceedings of 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Oct. 2016.
- [6] Alessandro Farinelli, Nicolo' Boscolo, Elena Zanotto, Enrico Pagello, "Advanced approaches for multi-robot coordination in logistic scenarios," *Robotics and Autonomous Systems*, pp. 34–44, Apr. 2017.
- [7] Karl Friston, "The free-energy principle: a unified brain theory?" *Nature*, pp. 127–138, Jan. 2010.
- [8] Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., Pezzulo, G., "Active Inference: A Process Theory," *Neural Computation*, pp. 1–49, Jan. 2017.
- [9] Ozan Catal, Tim Verbelen, Toon Van de Maele, Bart Dhoedt, Adam Safron, "Robot navigation as hierarchical active inference," *Neural Networks*, pp. 192–204, Oct. 2021.

- [10] David Eppstein, "Finding the k Shortest Paths," *SIAM Journal on Computing*, vol. 28, pp. 652–673, 1998.