

Title	サプライチェーンマネジメントへの強化学習の適用
Author(s)	田野, 勇二
Citation	
Issue Date	2000-03
Type	Thesis or Dissertation
Text version	author
URL	<a href="http://hdl.handle.net/10119/632">http://hdl.handle.net/10119/632</a>
Rights	
Description	Supervisor: 吉田 武稔, 知識科学研究科, 修士

修 士 論 文

サプライチェーンマネジメントへの強化学習の適用

指導教官 吉田 武稔 助教授

北陸先端科学技術大学院大学  
知識科学研究科知識社会システム学専攻

850055 田野 勇二

審査委員：吉田 武稔 助教授(主査)

中森 義輝 教授

本多 卓也 教授

2000年2月

# 目 次

第一章	はじめに .....	1
1.1	背景 .....	1
1.2	構成 .....	3
第二章	システム方法論.....	4
2.1	知識に関する考察.....	5
2.2	TOTAL SYSTEMS INTERVENTION.....	7
2.3	SCM への TSI 適用に関する考察.....	11
2.4	ソフト・システムズ方法論.....	12
2.4.1	論理的探索の流れ .....	12
2.4.2	文化的探索の流れ .....	14
2.5	まとめ.....	14
第三章	ビールゲームへの SSM の適用.....	16
3.1	ビールゲーム .....	16
3.2	SSM の適用.....	18
3.2.1	ビールゲームの論点.....	18
3.2.2	ビールゲームの関連システム.....	19
3.2.3	基本定義と概念モデル .....	20
3.3	シミュレーションの導入 .....	21
第四章	シミュレーションモデルの構築.....	23
4.1	シミュレーションの枠組み.....	23
4.1.1	顧客需要の変動.....	23
4.1.2	情報の時間遅れ.....	24

4.2	ヒューリスティック関数について.....	24
4.3	強化学習への定式化.....	25
4.3.1	強化学習の構成要素.....	26
4.3.2	Q値学習の設定.....	27
	(a) 状態の決定方法.....	28
	(b) 行為選択について.....	30
	(c) 報酬の設定.....	30
	(d) 変数の設定.....	31
4.4	まとめ.....	32
<b>第五章</b>	<b>シミュレーションの結果および考察.....</b>	<b>33</b>
5.1	ヒューリスティック関数の場合.....	33
5.2	強化学習の適用結果.....	36
5.2.1	情報遅れのない場合.....	36
5.2.2	情報遅れのある場合.....	37
5.3	シミュレーション結果の考察.....	43
5.3.1	プレイヤーごとの改善案.....	43
5.3.2	ビールゲームの論理的分析.....	44
5.3.3	ビールゲームの文化的分析.....	45
<b>第六章</b>	<b>終わりに.....</b>	<b>46</b>
6.1	総括.....	46
6.2	今後の課題.....	47
	謝辞.....	49
	参考文献.....	50

# 目 次

図 2-1	4つの知識変換モード.....	5
図 2-2	SSMにおける論理的・文化的分析の流れ.....	15
図 3-1	ビールゲームのアウトライン.....	17
図 3-2	基本定義のCATWOE.....	20
図 3-3	ビールゲームにおける論理的分析.....	22
図 4-1	シミュレーションの枠組み.....	24
図 4-2	強化学習の構成要素.....	26
図 4-3	Q値学習.....	27
図 4-4	強化学習の変数.....	28
図 4-5	状態の決定.....	29
図 4-6	行為選択の条件.....	30
図 5-1	ヒューリスティック関数の`在庫 - 受注残`履歴(情報遅れなし).....	35
図 5-2	ヒューリスティック関数の`在庫 - 受注残`履歴(情報遅れあり).....	35
図 5-3	強化学習を適用した場合の`在庫 - 受注残`履歴(情報遅れなし).....	39
図 5-4	`在庫 - 受注残`のサプライチェーン平均(情報遅れなし).....	39
図 5-5	各週のコスト小計(情報遅れなし).....	40
図 5-6	コストの履歴(RL:強化学習,Heu:ヒューリスティック, 情報遅れなし)...	40
図 5-7	強化学習を適用した場合の`在庫 - 受注残`履歴(情報遅れあり).....	41
図 5-8	`在庫 - 受注残`サプライチェーン平均(情報遅れあり).....	41

図 5-9 各週のコスト小計(情報遅れあり) .....	42
図 5-10 コストの履歴(RL:強化学習,Heu:ヒューリスティック, 情報遅れあり) ...	42

# 表 目 次

表 2-1	問題の内容によって分類されるシステム方法論のカテゴリー .....	8
表 5-1	サプライチェーンコスト結果(情報遅れなし) .....	33
表 5-2	サプライチェーンコスト結果(情報遅れあり) .....	34

# 第一章

## はじめに

### 1.1 背景

昨今、製造および部品・資材供給業者から顧客まで扱い、さらにグローバルオペレーションを含めたサプライチェーンを管理運営する手法として、Supply Chain Management (SCM)と呼ばれる経営手法が注目を集めている[1]。SCMは、サプライヤーから顧客へ向かう原材料や製品の流れ、物流に伴う情報の流れやキャッシュフロー、逆に顧客等の下流からサプライチェーンの上流工程への様々な要素を管理することで、サプライチェーン全体の収益向上・コストの削減・製品納期の短縮等を目指している。別の定義としては「不確定性の高い市場変化にサプライチェーン全体をアジル(機敏)に経営行動を対応させ、ダイナミックに最適化を図ること」としている。このようにSCMを行う際には、環境の変化を素早く反映するために、意思決定における迅速さが必要である。また今後、大規模なサプライチェーンが一層必要になることが予想される。

しかし、現状ではSCMは取り組みにくい問題として存在している。それは、様々な形態のサプライチェーンにそれぞれ、いくつかの短期的な目標が混在しており、更にそれらの目的は密接に関わっていることが多いためである。特に(1)原材料・製品の流通、(2)キャッシュフロー、(3)情報システム、(4)企業組織の協調作業について取り組まなくてはならない。(1)および(2)はサプライチェーンを流れる実体である。(3)は、(1)や(2)を扱うために、サプライチェーンを流れる情報を含めたシステムである。(4)はサプライチェーンの流れを決定する。特に、情報技術の進

展は、一般企業における大規模企業間取引システム等の、より広範な情報システム支援を可能にし、その他の問題を解決するために利用されているため、SCMを現実のものとしつつある。しかし、現状としてSCMを支援する情報システムは、取引処理に重点を置いたものに留まり、サプライチェーン全体を俯瞰した、適切な意思決定を行うための枠組みは、未だ発展途上である。なぜなら、SCMはさまざまな業務活動を含んでいるため、SCMを支援する情報システムは膨大な量の業務情報を取り扱わなくてはならない。その結果、指数関数的に増える実行可能な業務の組み合わせの中から、適切な選択を行うのは、大変困難だと考えられてきたからである。この例が示すように、問題を解決するために目的が不明瞭な状態で、闇雲にSCMを実行し、簡便な指標やコストのみで評価しても、サプライチェーンにとってパフォーマンスの向上には至らない場合が多い。逆に、サプライチェーン全体に悪影響を及ぼす可能性もある。これは、SCMの問題解決を目指す人が、SCMに意図的にかかわろうとする人間活動、および全体論的なパフォーマンス評価について十分考慮する必要があるのに、考慮していないためである。サプライチェーンという企業(グループ)文化に適したサプライチェーン構築を達成するためには、サプライチェーン内の協力関係を見通すこと、特に人間活動について理解することが重要である。

取り組みが困難とされるSCMの先行研究としては、サプライチェーンの基本モデルとしてSupply Chain Councilが公開したSupply Chain Operations Reference - model (SCOR)がある。SCORを参考とすることにより、サプライチェーンの構成要素を定義し、サプライチェーンの構成からベンチマーキングによる評価までの一連の手続きを手法としたモデリング理論がある[2]。SCORは企業間ロジスティクスにとって、サプライチェーンプロセス類型による、共通の言葉を用いたSCMの認識を可能とし、サプライチェーンにおける競争上の目標が明確にした。さらに、ベンチマーキングを用いることで、企業内外の優れた業務手順(ベストプラクティス)と比較することにより、業務の改善を行うプロセスを組み込んでいる。SCORの方法論は、サプライチェーンのパフォーマンス評価を含んでいるため、有効な解決法の一つである。しかし、この手法はすでにすべてのプロセスが判明した状態では有効であるが、何を改善すればよいかわからない状態では、適用することが難しい。

本研究では、環境の変化に対して柔軟に対応できる SCM 意思決定支援システムを、シミュレーションを通じて考察する。そのために、Total Systems Intervention (TSI) [3]とソフト・システムズ方法論(SSM)[4][5]と呼ばれるシステム方法論を用いることで、SCM を行うために必要な情報を、人間活動システムとして明確に含めたモデル構築によって検討する。ここで、サプライチェーンの意思決定を支援するアルゴリズムとして、強化学習[6]の枠組みを定式化した。

## 1.2 構成

本論の構成について説明する。第二章では、TSI および SSM を用いて SCM を考察し、サプライチェーンにとって有効な改善方法を発見する手続きについて説明する。第三章では、SCM に対して SSM を適用し、サプライチェーンに対する改善案を模索する。本研究では、一般的なサプライチェーンの事例からではなくビールゲームと呼ばれる、ロールプレイングシミュレーションを用いて説明する[7][8]。この章で構築された概念モデルは、第四章で示す強化学習のシミュレーションモデルの構築に活用する。第四章では、第三章で求めた概念モデルが示すサプライチェーンの概念を含めたシミュレーションの枠組みを示す。シミュレーションは、知覚される現実の世界との比較を行うために用いる。さらにシミュレーションでは、強化学習の枠組みにより、各プレイヤーの意思決定を支援させる。第五章では、ヒューリスティック関数および強化学習を適用したビールゲームのシミュレーションの結果から、サプライチェーンに求められる行為が強化学習によってどのように表されるかについて述べる。この結論から、ビールゲームのサプライチェーンに提案できる有効な改善案を提示する。第六章において、TSI・SSM および強化学習を用いたシミュレーション結果から、TSI や SSM が SCM にとって有効なシステム方法論であること、ならびに強化学習を用いたシミュレーションを、システム方法論における学習システムの中に組み込むことで、サプライチェーンに有効な改善案が提示できることを示す。

## 第二章

# システム方法論

サプライチェーン構築においては、サプライヤーから顧客へ向かう原材料や製品の流れ、それらに伴う情報の流れやキャッシュフロー、それらの逆の流れに関するすべての活動を考える必要がある。さらに、これらに付随する意図的な人間活動についても十分考慮する必要がある。特に、このような人間活動を考慮することが、理想的な(または最適な)サプライチェーン構築を達成するために重要となる。

しかし現状では、通常最適化技術やソフトウェア開発プロジェクトなどにこのような人間活動まで含めて考えることは困難な作業となっている。このため、人間活動に関する問題を考察の範囲外として無視し、意図的に排除してしまうことが多い。このようなことが原因で、企業文化や風土になじまない、また変化に対応できないソフトウェアシステムが構築されてしまう結果となる。

本章では、人間活動を積極的に考えたシステム方法論である Total Systems Intervention(TSI)[3]およびソフト・システムズ方法論(SSM)[5]によって SCM に必要な方法論とその適用のために必要な手続きを考察する。これらのシステム方法論は、相互に関連する人間活動の集合またはこのような意図的な人間活動を行っている人間の集合を「人間活動システム(Human Activity System)」と呼び、その中心に位置づけている。

まず人間活動システムに密接に関連している知識創造について考察し、次に TSI および SSM について概説する。この際、これらのシステム方法論が、知識創造プロセスを実現するためのひとつのアプローチとして適切であることを意識した概説を試みる。

## 2.1 知識に関する考察

本節では、人間活動システムに関するシステムアプローチについて考察する。この人間活動システムの本質は、人間の知識創造活動またはプロセス、しいては知識創造活動を実施する人間そのものと等価である。本節では、これらのことについて、文献[9]に沿って考察する。

例えば、人間活動を通して、ビジネス・ロジスティクス概念が創造され、現在サプライチェーンへと進化している。また、連続的イノベーション概念により、サプライチェーンはビジネス・ロジスティクスの成功を引き継いでいると説明可能であろう。これらの成功要因のひとつはメタファーおよびアナロジーの採用である。これらは、常に不明瞭さまたは冗長さを持っている。しかしながら、このような性質が、形式的に説明することが困難な問題を理解するのを助けている。それがすなわちイノベーションへと導くのである。

このような考察における次元として、知識創造における認識論的次元がある。ここでは、暗黙知と形式知の社会的相互作用が知識創造の鍵であり、イノベーションの源であると主張されている。さらにこの2つの知の社会的相互作用を通じて知識が創造される事を前提に4つの知識変換モードが提案されている(図2-1)。



図 2-1 4つの知識変換モード

ここで、表出化(Externalization)とは、暗黙知を明確なコンセプトに表すプロセスである。これは、暗黙知がメタファーなどを通じて次第に形式知として明示的になっていく点で、知識創造プロセスの真髄であるものとされている。さらに連結化(Combination)とは、コンセプトを組み合わせてひとつの知識体系を創り出すプロセスとして提案されている。すなわち、個々のコンセプトだけではなく、全体をひとつのコンセプトとしてみるシステム的な知識の重要性も主張されている。

個々人の知識および価値観の共有も重要である。例えば、個々人の知識および価値観は、メンタルモデルの形で与えられるという考え方がある[10]。ただし、このようなメンタルモデルは、人間の内部に存在するモデルとして考えられている。また、このような個々人の知識を共有する仕組みのひとつとしてシステムダイナミックスの導入が提案されている[11]。そこでは一種のシミュレーションモデルを積極的に利用することにより、知識や価値観の共有、さらに組織学習へと導こうとしている。

しかし、メンタルモデルの幾分かは知識創造の変換モードにおける表出化により、概念として形式化することが可能であるものと考えられる。この考え方に沿って、SSM[5]を採用した。

SSM はシステム思考の考え方に基づいて、概念モデルを導出し、その概念モデルに基づいたディベートにより個々人の知識および価値観を共有するようなシステムアプローチとして提案されている。さらに概念モデルすなわち人間活動システムとして、創発特性、階層構造、およびコミュニケーションとコントロールのプロセスを持ち、変化する環境の中で原理的には生存できる全体的抽象的な概念に基礎をおいている。

このように SSM では、表出化における概念知識としての概念モデルが主要な役割を果たす。さらに連結化におけるシステム全体の知識(システム的な知識)の考え方も含まれている。このようなことを背景として、概念モデルの創造および概念モデルに基づいたディベートの実施は、表出化および連結化のモードまたはプロセスの具現化として、位置づけられるものと想定する。ここで、このようなシステムアプローチの適用に妥当性を与えるメタシステム方法論として、TSI が提案されている。

## 2.2 Total Systems Intervention

Total Systems Intervention(TSI)[3]とは人間活動システムに密接に関連した問題解決法であり、知識創造(Creativity)、選択(Choice)、実装(Implementation)の3つのフェーズから構成されている。これらの特徴は以下の通りである。

### 知識創造(Creativity) :

- ・ タスク 目標、考慮点および問題を浮き彫りにする
- ・ ツール メタファー
- ・ 成果物 主たる問題を浮き彫りにしている支配的なメタファーおよび従属的なメタファー

### 選択(Choice) :

- ・ タスク 適切なシステム方法論の選択
- ・ ツール システム方法論システム(表2 - 2)。これはメタファーとシステム方法論の関連を明確にしている
- ・ 成果物 適用可能な支配的なシステム方法論および従属的なシステム方法論

### 実装(Implementation) :

- ・ タスク 改革のための提言およびそれらの実装
- ・ ツール 選択されたシステム方法論
- ・ 成果物 適切で調和のとれた改革

次に各フェーズの説明を示す。

知識創造フェーズ；

問題解決手順として、まず知識創造フェーズでは、主に次の5つのメタファーを採用している。

- ・ 機械的メタファー(machine metaphor)または閉じたシステムビュー(closed system view)
- ・ 生体的メタファー(organic metaphor)または開いたシステムビュー(open system view)
- ・ ニューロ・サイバネティクスメタファー(neuro-cybernetics metaphor)または生存可能なシステムビュー(viable system view)
- ・ 文化的メタファー(cultural metaphor)(社会的規範や価値観を強調)
- ・ 政治的メタファー(political metaphor)(例えば、単一政府制のようなチーム、複数政府制のような連立、独裁制のような刑務所)

表 2-1 問題の内容によって分類されるシステム方法論のカテゴリ-[3]

支持者 システム	単一政党支持	複数政党支持	強圧制
単純	<ul style="list-style-type: none"> <li>・オペレーションズリサーチ</li> <li>・システム分析</li> <li>・システム工学</li> <li>・システム・ダイナミクス</li> </ul>	<ul style="list-style-type: none"> <li>・社会組織構築</li> <li>・戦略的な仮説検証 -定量評価 -仮説評価</li> <li>・論証法としての討論</li> <li>・総合法</li> </ul>	<ul style="list-style-type: none"> <li>・批判的なシステム・ヒューリスティクス</li> </ul>
複雑	<ul style="list-style-type: none"> <li>・生存可能システム診断</li> <li>・一般システム理論</li> <li>・社会科学システム思考</li> <li>・偶発事象理論</li> </ul>	<ul style="list-style-type: none"> <li>・対話型計画法</li> <li>・ソフト・システムズ方法論 (SSM)</li> </ul>	現在見つからない

これらのメタファーが、対象となる組織の性質に応じて繰り返し適用され、適用ごとに支配的なメタファーおよび従属的なメタファーを認識しながら、適切な問題の理解が得られる。このような適切な理解を得る仕組みを適用しない場合、例えば機械的メタファーだけの使用は、仕様が決定され、閉じたシステムビューだけを利用するような大規模な情報システム開発において見られ、数々の失敗例の原因となっている。よって、ここで述べられている5つのメタファーを種々の観点から適用する必要がある。

### 選択フェーズ：

選択フェーズは、まず与えられた問題を次の2つの次元で区別することから始まる。

- ・ 問題に関与する人々の利害や価値観などに着目(関与者の次元と呼ぶ)
- ・ 問題の複雑さに着目(システム次元と呼ぶ)

システムの次元には、問題の内容に応じて主に2つの区分がある。ひとつは「単純なシステム(simple system)」であり、もうひとつは「複雑なシステム(complex system)」である。さらに、単純なシステムの特徴として次の項目を挙げている。

- ・ 構成要素のすべての属性は既知である。
- ・ 構成要素間の相互干渉には規則性がある。
- ・ 振る舞いには規則性がある。
- ・ システムは静的である。
- ・ サブシステムは目標を追求しない。
- ・ システムは振る舞いに影響を与えるものからの影響を受けない。
- ・ システムは閉じている。

複雑なシステムは次のように特徴づけられている。

- ・ 構成要素の属性は既知でない。
- ・ 構成要素間の相互干渉には規則性はない。
- ・ 振る舞いは確率的である。
- ・ システムは動的である。

- ・ サブシステムは意図的に振る舞い、自身の目標を生成する。
- ・ システムは、振る舞いに影響を与えるものからの影響を受ける。
- ・ システムは開いている。

次に関与者の次元としては、政治的なメタファーを用いた単一政府制支持者(Unitary)、複数政府制支持者(Pluralist)および強圧政府制支持者(Coercive)の3つの区分を設定している。これらは、主に問題に関与している人々または組織の利害関係を元にした同意または不同意などの関係を表している。これらの特徴は次の通りである。

単一政府制支持者(Unitary)：

- ・ 興味を共有している。
- ・ 信念および価値観は一致している。
- ・ 目的と手段に合意している。
- ・ 全員が意思決定に関係する。
- ・ 合意された目的に従って行動する。

複数政府制支持者(Pluralist)：

- ・ 興味は基本的には矛盾していない。
- ・ 信念および価値観にはある程度広がりがある。
- ・ 目的と手段は必ずしも合意されていないが、それらに関する妥協はありえる。
- ・ 全員が意思決定に関係する。
- ・ 合意された目的に従って行動する。

強圧政府制支持者(Coercive)：

- ・ 共通の興味を持たない。
- ・ 信念および価値観はまちまちである。
- ・ 目的と手段に合意はなく、妥協もない。
- ・ 目的と手段に合意はなく、妥協もない。
- ・ 意思決定において、ある特定の人(または人々)が他を圧倒した決定をおこなう。
- ・ 目的に関する合意は不可能である。

このような次元の定義にそって問題を類別し、さらに知識創造フェーズに示すようなメタファーを参考にして、問題を的確に表現する支配的なメタファーおよび従属的なメタファーを明確にする。例えば、すべての問題に対して、他で用意され現状に適しないマニュアルを用い、その記述を規範として実行する場合には、しばしば強圧政府制支持者のものになるものと想定できる。

このようにして、システムの複雑さおよび関係者の次元が決定すると、表 2.1 にそって適切なシステム方法論が選択可能となる。ここでは、知識創造フェーズで明確にされた支配的および従属的なメタファーに対応したシステム方法論を選択することになる。

#### 実装フェーズ：

実装フェーズでは、選択フェーズで選択されたシステム方法論を用いることにより問題解決を図る。

このような3つの問題解決フェーズは繰り返され、その都度、支配的および従属的なメタファーや、それらに対応したシステム方法論も変化していくことになる。

## 2.3 SCM への TSI 適用に関する考察

サプライチェーンマネジメント(SCM)に TSI における2つの次元の定義を適用すると、それは複雑なシステムで複数政府制支持者としてのメタファーのカテゴリに属するものと認識できる。よって、表 2.1 に示す適用可能なシステム方法論から、SSM を採用し、その本質に迫る。ここで、SCM の構造は、その環境に応じて急激に変化することに注意する必要がある。よって各要素に着目した世界観より、全体をひとつとして捉えるシステムミックな世界観を持ち、その世界観に基づいたビジネスプロセスを構築することが重要である。このような意味において、システムミック

な世界観を重要視し、概念モデルにより知識および価値観を共有または体系化する、TSI および SSM の連携には妥当性があるものと想定する。

## 2.4 ソフト・システムズ方法論

ソフト・システムズ方法論(SSM)は、システム思考を基礎として、人間の価値観や世界観からくる多様性をシステム方法論として扱い、それらを共有することによって問題の抜本的な改善を生み出すための方法論である。ここで、システムとは、個々の要素の集合であり、これらの要素は、この集合がひとつの個別体としての特性をもつ全体を構成するように相互に結びついているものと定義されている。

SSM の適用の狙いは、何らかの意味で不満足な状況の改革にあり、そのような改革は「システムの的に望ましく (systemically desirable)、文化的に実行可能(culturally feasible)」でなければならない。また、この手法で主たる役割を果たす「概念モデル(conceptual model)」は、「反論可能(defensible)」か「反論不可能(indefensible)」か、といった知覚された現実世界に関するディベートを行うのに適切かどうかだけが問われる。

SSM では、現実世界の状況に関連がありそうなモデル(注：現実を表すモデルではなく、システム思考における概念のモデル)をいくつか作り、比較を行うことによって、それらのモデルを現実状況の知覚と照らし合わせることをその本質とする [12]。この比較を行うことによって、吟味している現実世界に対する意図的行為を決定するディベートが導かれる。このプロセスは、「論理的探索の流れ」と「文化的探索の流れ」およびそれらの相互作用によって構成されている。SSM に関する分析の流れを、図 2-2 に示す。次の項から、SSM における分析の流れを、特に論理的な分析に絞って説明する。これは、論理的分析は本研究の目的であるシミュレーションの位置付けを明確にするためである。

### 2.4.1 論理的探索の流れ

SSM の適用プロセスは、7 ステージ・モデルと呼ばれる。それらの主な流れは、

現実世界において、問題を発見し、システム思考により関連する概念的活動を明確にし、それらを現実世界と比較することにより、改革を実行するフェーズを繰り返し実施することである。さらに、これらを随時繰り返すことである。

まず「基本定義」を明確にする必要がある。基本定義とは、システムの所有者(自分の意向に沿わなければ、このシステムの活動を停止できる人)の長期的な目標 Z を達成するために、ある方法 Y によって、ある特定の変換プロセス X を実施するシステムを定義している。SSM では、現実世界の基本定義を「人間活動システム」として構築する。すなわち、創発特性、階層構造、およびコミュニケーションとコントロールのプロセスをもち、変化する環境の中で原理的には生存できる全体的抽象的な概念と定義している。

以下では、このアプローチの用語により、これを「目標」と呼び、「長期的目標」を「目的」と呼ぶことにする。システムの創発性の観点からは目的をもつことは重要であるが、それらの目的は一般に達成が困難であり、その評価も難しい。

次に基本定義の要件を満たすために、最小限必要な活動を洗い出し、これらの活動を論理依存性に基づいて図示する。さらに原理的には生存可能であることを保証するために、モニタリングとコントロールの2つのプロセスを付加する。

さらに「変換」の概念には、その入力から出力への変換の成功を見る必要がある。このために、可動性(この手段でうまくいくのか)、効率性(どのくらいの資源を利用したか)、有効性(目標に合致しているか)の3つの基準を決定しておくことが必要となる。

以上により、SSM でいう「反論可能な概念モデル」に相当する概念モデルが導出できる。本研究ではさらに、SCM を行うために必要なモデルは、目的を達成するための総合的な人間活動の一部として作成されるべきであると主張する。よってこれまでに作成した概念モデルに対して、さらにモニタリングとコントロールのプロセスを考え、そこに有効性(目的に合致しているか)の基準を明確にしておく。

これにより導出された概念モデルと「知覚された現実」とを比較することにより、「調和」や「妥協」の状態としての、アコモデーションを見つけ出すプロセスを実施する。このためには企業文化や風土といったものを取り扱う文化的探索が必要となる。

## 2.4.2 文化的探索の流れ

文化的探索には、第1分析、第2分析、第3分析の3つの流れがある。第1分析は「介入の分析」と呼ばれ、依頼者や所有者などを明確にする役割分析を実施する。第2分析では、インタビューなどの結果に応じて、役割、規範および価値の3つの面について繰り返し分析する。第3分析では、人間を取り巻く政治的要因を慎重に調査分析する。

以上の論理的および文化的探索により、作成者間でのコミュニケーションおよびコラボレーションが達成でき、サプライチェーンの所有者や、ソフトウェア構築に携わる人にとって有効な考察手法となりうる。

## 2.5 まとめ

この章では、Total Systems Intervention と呼ばれるメタシステム方法論から、サプライチェーンを考察する適切な方法論としてSSMが導かれた。さらに、SSMによって、サプライチェーンに存在する問題状況を改善しようとする人が、企業文化に適した望ましい改善点を提示することができる枠組みを提示することができた。

このように、問題を解決しようとする人々にとって、問題状況について概念モデルを用いて共有することができれば、SSMの比較のステージにおいて、現実世界に改善案を示す。

次章より、ビールの流通過程であるビールゲームについて、強化学習を用いた離散事象シミュレーションによって試行・評価する。ビールゲームの基本定義から概念モデルを構築し、シミュレーションに有効に利用する。

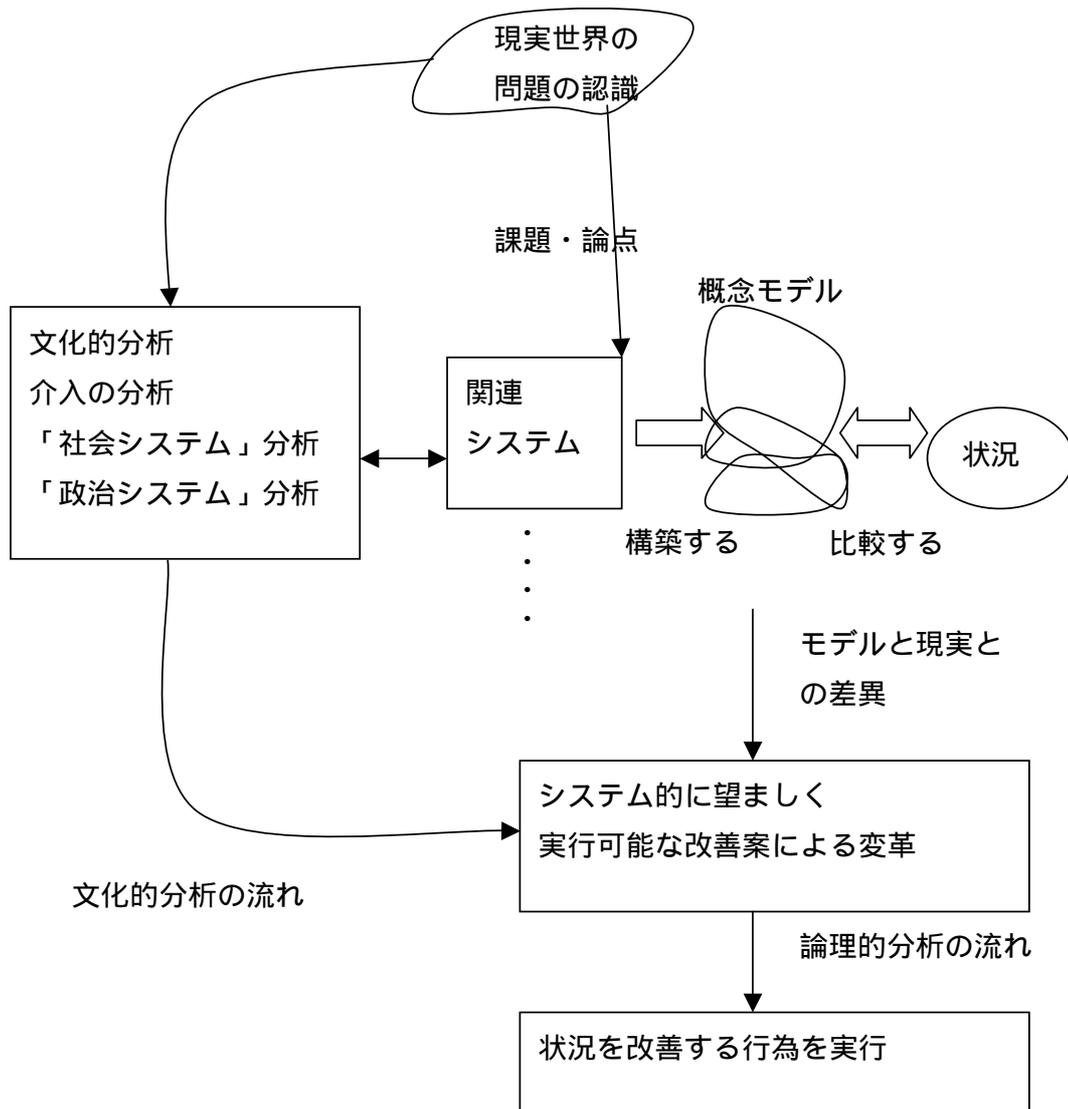


図 2-2 SSM における論理的・文化的分析の流れ

## 第三章

# ビールゲームへの SSM の適用

本章では、ビールゲーム[7][8]の問題を SCM の一例として取り上げる。ビールゲームは流通過程をめぐるロールプレイングシミュレーションであり、SCM を在庫管理と意思決定の観点から理解するひとつの題材として捉えられる。

まず、全般的なビールゲームの内容について述べる。次に、ビールゲームの問題定義を、SSM に基づいて、ビールゲームに存在する問題、目的、行動として関連システムで表す。さらに関連システムから基本定義を作成し、ビールゲームを概念モデルとして示す。

### 3.1 ビールゲーム

ビールゲームは、MIT で行われたゲーム理論に関する研究から始まったシミュレーションロールプレイングゲームである[7][8]。以来、多くの学生や教授らが参加し、景気の循環作用およびその複雑性について研究がなされた。本節ではこのビールゲームについて説明を述べる。

ビールゲームでは、そのビールを取り扱う業者に各プレイヤーが成り代わり、顧客の注文量がある週に増加した場合の影響の増幅状況を調べることができる。このシミュレーションは、発注量を決めるための意思決定に、さまざまな指標が必要となるため、判断の難しい SCM である。図 3 - 1 に、ビールゲームの簡単な手順を示す。

まず、このゲームは 4 組の業者(Retailer=小売業, Wholesaler=卸業, Distributor=配送業, Factory=工場)を行為主体(プレイヤー)が担当する。毎週、顧客はビールをケ

ース単位で小売商に要求する。小売商は要求されたビールを出荷する。このときビールは在庫から引かれる。小売商の番になると、ビールを卸売りに注文する。卸は自分の在庫から要求された量を出荷する。同様に卸は、配送業者に注文し出荷してもらう。配送業者の番になると次は製造に注文し出荷してもらう。製造ではビールを生産する。どのステージでも、出荷遅れと注文受け付け遅れが発生する。これらは毎週ごとに進められる。ゲーム中は顧客の注文量は伏せられている。つまり、小売以外の各プレイヤーは隣り合ったプレイヤーの在庫状況のみを知ることができる。この設定に基づいて、ゲームは進められる。

ビールゲームはコストの最小化を求めるゲームである。ここでコストとは

- ・ 0.50\$/case/week(在庫所有に伴う経費)
- ・ 1.00\$/case/week(受注残務が残っている場合に発生する経費)

とされる。最後に顧客の購買情報として5週間後にこれまでの週4ケースの注文が週8ケースの注文になることとする。この場合各行為主体は、在庫を最小化したいが、在庫切れによる納期遅れは避けたい思惑があるために調整が難しい問題である。

#### 手順

- 1 : 商品を受け取ってから製品を配送する。先週送られた製品は在庫に追加される。
- 2 : 注文量を発送する。満たすべき発送量は注文分+先週からの残務分である。注文は在庫の許す範囲内だけ満たされる。もし満たされなかった注文分がある場合は受注残とし記録される。
- 3 : 在庫および先週からの残務を記録する:
- 4 : 顧客の注文を受け取り、発注量が伝えられる工場は製造要請分だけ製造遅れに移動する。
- 5 : 注文提出:各プレイヤーは発注量を決定し伏せておく

図 3-1 ビールゲームのアウトライン

## 3.2 SSM の適用

本節では、ビールゲームに対する SSM の適用手続きを示す。

まず、ビールゲームのいくつかの論点から、関連システムの一つを導き、その基本定義および概念モデルを求める。本節で求めた概念モデルと現実世界に関する知覚とを比較することで、ビールゲームに対してどのような変革を求めているかが明らかになる。

### 3.2.1 ビールゲームの論点

ここでビールゲームの論点を、3.1 節で示したビールゲームの問題定義を元にして文化的・論理的に分析する。

第一に、サプライチェーンにおいて、協働作業を行っていないことが挙げられる。ビールゲームでは、各プレイヤーは他のプレイヤーに関する情報を、あまり重視しないために、自己中心の意思決定を行っていると想定した。

次に在庫と受注残を抑える必要がある。なぜなら在庫が多いことにより、在庫の保管が必要となるため無駄が多くなる可能性があるためである。しかし、在庫を少なくすることにより受注残が増える可能性がある。受注残は、顧客のほしい商品が、購買時期に購入できないために発生する。この受注残は納期遅れを引き起こし、ひいては将来的な需要を減らす可能性がある。在庫と受注残は、ビールゲームの SCM ではトレードオフの関係にあり、どちらも重要である。

さらに、情報遅れについて考察する。現実のサプライチェーンでは情報システムの改善が、サプライチェーンの価値向上に役立っているため、情報遅れによる影響を考察することは有効である。

また、顧客満足度を設定する必要がある。しかし、通常であれば、顧客満足度をアンケートに基づくデータによって調査することが想定できる。ビールゲームにおいては、顧客が購入できないことがビールゲームのプレイヤーにとって機会損失であると同時に、受注残を増すために顧客満足度を低下させる原因になる。

ビールゲームでは、顧客の一時的な需要変化が、上流に流れることで増幅される、いわゆる変動の拡大(あるいはブルウィップ)と呼ばれる現象が発生する。そこで、このビールゲームのサプライチェーンでは、変動の拡大について、金・物・情報の移動に伴ういくつかの指標から、導かれた改善案に対して可動性・効率性・有効性が認められるかを示すことが肝心である。

### 3.2.2 ビールゲームの関連システム

前述したように、ビールゲームに複数の論点がある。先に示した論点から、それぞれを関連システムとして述べる。

まず、各プレイヤーにとってシステムとは

- ・ 上流および下流からの製品情報から、意思決定を行うシステム
  - ・ コストを最小化するシステム
- である。次に、サプライチェーン全体では、
- ・ 顧客満足度を向上させるシステム
  - ・ 複数企業が協働作業を行うシステム
  - ・ サプライチェーン全体のコストを抑えるシステム
  - ・ 継続したサプライチェーンの連携を目指すシステム

といえる。

複数のプレイヤーの意識が SCM を行うことでメリットがあると認識することができれば、持続したサプライチェーンが望める。

サプライチェーン全体と各プレイヤーの関係は、親と子の関係に見える。しかし、ビールゲームが示すサプライチェーンの性質として、一部の滞りは全体に波及することが挙げられる。このことから、2者の関係は、4輪の車というメタファーによって表せる。サプライチェーンという車を動かすには小売・卸・配送・工場といったタイヤの存在すると共に、それぞれが連携した動きをしないと、まっすぐに速く進めない、という意味をもつ。それは、各プレイヤーの目標の達成は、サプライチェーン全体の目標の達成に他ならないことを示している。

### 3.2.3 基本定義と概念モデル

先に述べたビールゲームのシステムにおいて求められる目標およびメタファーから、ビールゲームのサプライチェーンを表す基本定義を、

“顧客満足度の高いサプライチェーンを構築するために、協働作業を行い、受注残を解消しつつ、在庫コストを抑える SCM を行うシステム”

と定義づけた。この定義の主たる活動は「SCM を行う」ことである。達成すべき長期的な目標は「顧客満足度の高いサプライチェーン」である。そのために受注残と在庫コストに焦点を置く。この基本定義に関する CATWOE を図 3 - 2 に、概念モデルを図 3 - 3 に示す。

CATWOE から、このシステムは既存の情報を共有しないサプライチェーンを、情報を共有するサプライチェーンに変換する。また、このシステムで変換されたサプライチェーンは、高い顧客満足度と低コストを維持できるという確信を持っている。

C(受容者)	サプライチェーンの全プレイヤー
A(行為者)	サプライチェーンの全プレイヤー
T(変換)	プレイヤーは独自の情報を用いて、個々のコストを抑えるように努めるサプライチェーン プレイヤー間で公開されている情報を元に全体のコストを抑えるように努めるサプライチェーン
W(世界観)	情報を共有したサプライチェーンは、高い顧客満足度と低コストを持続できるという信念
O(所有者)	企業グループの責任者
E(環境)	顧客需要の変動

図 3-2 基本定義の CATWOE

概念モデルは、基本定義の変換を表す人間活動プロセスが、論理的に依存した状態で表現された。この概念モデルは、協働作業を行えるようになったことで、サプライチェーンにおける各プレイヤーは在庫や受注量・受注残といったローカル情報と、別プレイヤーの在庫量などの様々な情報を手に入れられるようになった。この情報によって、発注量を決定することによりビールゲームのSCMが行われる。

概念モデルからSCMの変換が行われたことを判断するために、可動性・効率性および有効性についてモニタリングすることが必要である。ここで可動性とは「その手段がうまくいくか」をあらわし、効率性は「アウトプットを一単位出力するためにどれだけの資源を用いたか」を示す。最後に有効性は「より高いレベルの長期の狙いに合致しているか」を見るものである。この中で、可動性・効率性と有効性は短期的性質と長期的性質を判断する基準として分類される。特に、ビールゲームは経済活動であるので、可動性および効率性は経済性(コスト)として同義となり、情報の共有によって、どれだけ向上したかを評価する。また、有効性を示すには、サプライチェーン全体の評価として、高い顧客満足度を持続できるかを示す事とする。これらの基準が満たされるかどうかをモニタリングし、その評価に応じた行動としてモデルの変更を行える。

### 3.3 シミュレーションの導入

本章で示されたビールゲームの基本定義および概念モデルは、ビールゲームの認識を明確にした。さらに論理的分析を進めるために、現実のビールゲームに対するシステムの望ましく、文化的に実行可能な改善案を提示することが必要である。そこでサプライチェーンを離散事象シミュレーションによって表し、その振る舞いから改善案を示す。

第四章では、本章で示した概念モデルを満たすビールゲームのシミュレーションモデルを構築する。特に、各プレイヤーに意思決定を行うアルゴリズムとしてヒューリスティック関数および強化学習を適用する。これらのアルゴリズムによって選択された行為から、SCMに対する改善案を示す。

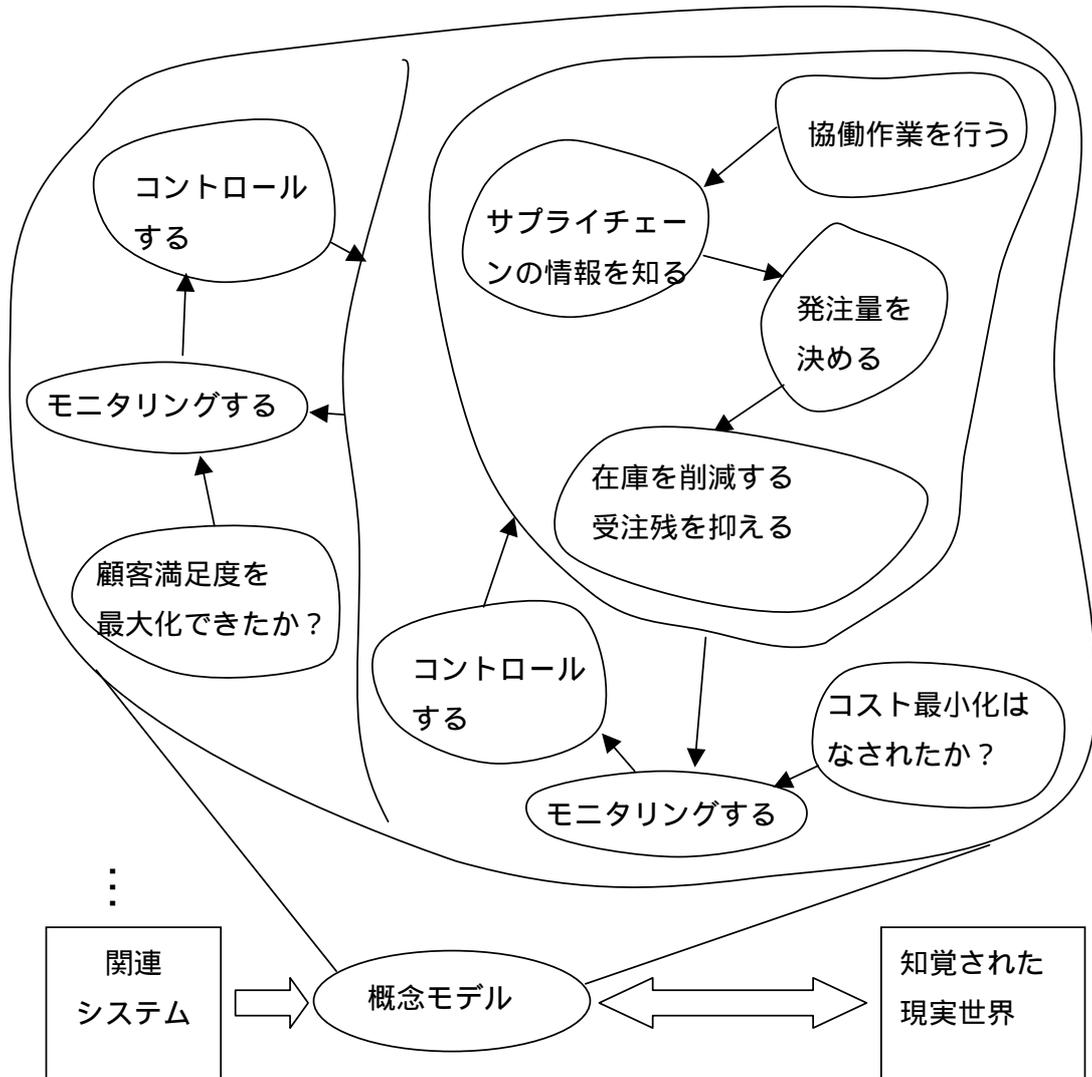


図 3-3 ビールゲームにおける論理的分析

## 第四章

# シミュレーションモデルの構築

本章では、第三章で求めたビールゲームの概念モデルが示す、ビールゲームが求める変革が行われる離散事象シミュレーションモデルを構築し、サプライチェーンの一現象である変動の拡大について取り組む。利用したアプリケーションは、Excel、VBA である。

本シミュレーションでは、ヒューリスティック関数を、協働作業を行っていないサプライチェーンにおける各プレイヤーの意思決定支援ツールとみなす。一方、協働作業を行うことにより、サプライチェーン全体の情報を知ることができる。強化学習の枠組みでシミュレーションを行うことにより、サプライチェーンが協働作業を行うことでどのようなサプライチェーンに変革できるか示す。

### 4.1 シミュレーションの枠組み

#### 4.1.2 顧客需要の変動

今回のビールゲームシミュレーションでは、顧客需要を毎週 4 ケースから五週間後に 8 ケースに倍増させる。この変化は、一般的な離散事象の中の一事象である。しかし、顧客の一時的な変化は、サプライチェーンの性質である「変動の拡大」を如実に示している。このことから、ビールゲームのシミュレーションによって、顧客需要の変動を抑えるために検証できる。三章で示したビールゲームのアウトラインに従ってビールゲームのプログラムを概略として示す。

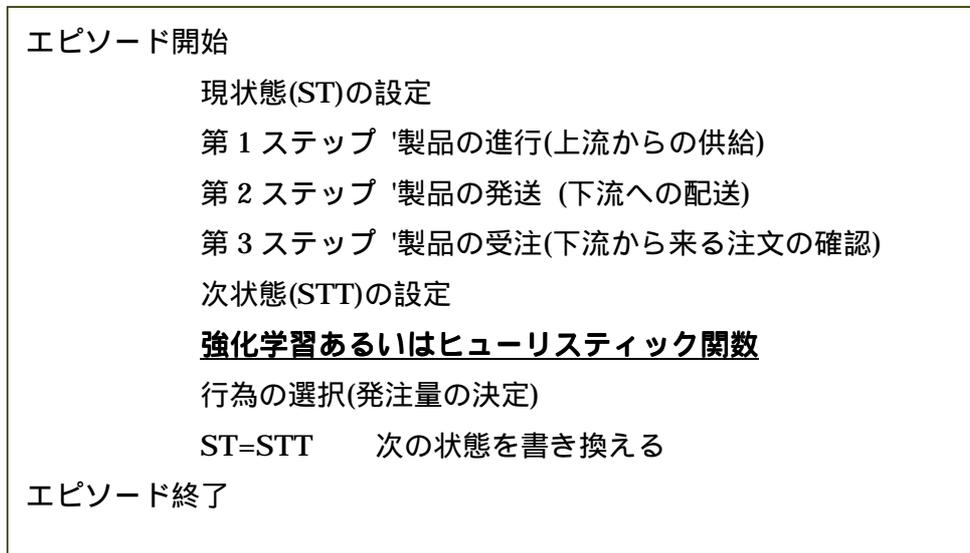


図 4-1 シミュレーションの枠組み

### 4.1.2 情報の時間遅れ

本研究のシミュレーションでは、顧客から工場にいたる発注量について、情報遅れがない場合とある場合に振り分けた。情報遅れがなければ、下流の発注量に関する情報は、次の週には上流の行程に到達する。しかし、情報遅れがある場合、下流の発注量に関する情報は、一時間単位(一週間)遅れて上流の行程に到達する。離散事象シミュレーションであるので、サンプリングする時間の単位は1時間単位刻み(一週間)とする。100時間単位を一区切り(エピソード)とし、その時点の状態を終端状態とする。

## 4.2 ヒューリスティック関数について

経験が必要な発注量の判断を試行錯誤的に決定した関数としてヒューリスティック関数がある。この関数にサプライチェーンのいくつかの情報をインプットするこ

とで、適切な発注がなされたとする。

まずヒューリスティック関数の目的として、

- ・ 在庫から予想される損失を差し引く
- ・ 現実の在庫と期待される在庫量との差を埋める
- ・ 残務のための十分な供給を維持する
- ・ 等があげられる。

まず情報としてプレイヤーに届く注文から今後の消費を予測し、次に理想的な在庫量を 8 ケースとして各自設定する。その差を必要な注文量に置き直す。さらに、理想的な供給量を 4 ケースとして設定し、現在の供給との差から注文を行う。それぞれ理想的な量より多い場合は注文しないものとする。これにより、毎週の発注量を決定する。

発注量=受注量 + 理想在庫量からの差 + 理想供給量との差

さらに、この観点から参考文献[7]から、発注量は負の値にはならない。そのため、決定した発注量が負、あるいは発注量が現在の在庫量の 1/4 以下ならば発注を行わないものとする。

このヒューリスティック関数は、強化学習の枠組みで得られた方策による発注量決定との比較のために構成した。本来なら、各プレイヤーに異なる判断基準があるべきだが、今回はすべてのプレイヤーを同一のヒューリスティック関数で意思決定を補う。

## 4.3 強化学習への定式化

この節では、離散事象シミュレーションとして表したビールゲームの SCM において、各プレイヤーの意思決定を支援するために強化学習[6]の概念を示す。強化学習は教師なし学習と呼ばれる。まず、強化学習の構成要素について述べ、強化学習にとって重要な価値関数の更新について示す。

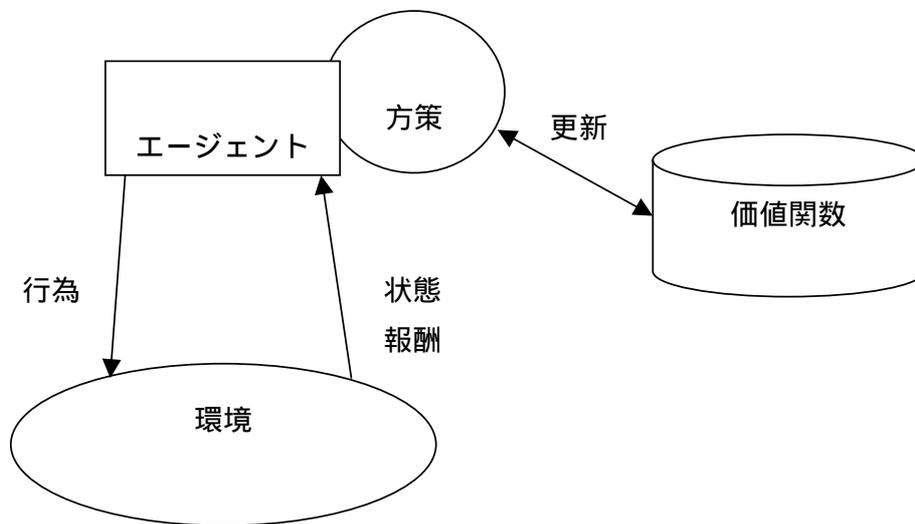


図 4-2 強化学習の構成要素

### 4.3.1 強化学習の構成要素

強化学習を構成する、いくつかの要素について説明する(図 4 - 1)。強化学習は、教師付の学習(これは教師が与えた正しい例を基に正しい出力を導く学習である)とは異なり、適切な回答を教えてくれる教師のいない環境における学習である。つまり教師なし学習にあたる。

強化学習では、行為を行う「エージェント」、エージェントが携わる「環境」、エージェントがある行動を起こした場合、それに基づく環境の変化から得られる「強化信号(報酬あるいは罰)」、得られた強化信号を元にさらに誤差逆伝播と呼ばれる手法により構築する「価値関数」があげられる。この「価値関数」はエージェントの「方策」と呼ばれる意思決定モデルと密接に関係しており、環境の状態からエージェントの行為を決定する。任意のゴールにたどり着くまでの試行(エピソードと呼ぶ)を複数回行うことで価値関数を更新し適切な方策を得ることを「条件付け」と呼んでいる。また価値関数とは、ある状態に到達したことによって将来どの程度報酬が得られるかを表す期待値のことである。環境の状態のみで決定する場合を「状態

価値関数」と呼び  $V(s)$  と表す。また状態と行動の組み合わせによって得られる期待値を指す場合もある。これを行為価値関数と呼び  $Q(s, a)$  と表す。ここで  $s$  は状態、 $a$  は行動を意味する。

### 4.3.2 Q 値学習の設定

本章において、シミュレーションに利用した強化学習の枠組みは、時間差分学習法のひとつである Q 値学習を用いている。Q 値学習は、C.J.C.H.Watkins によって考案された手法である。そのアルゴリズムを示す(図 4 - 2)。

図 4 - 2 内の式(1)が示すように、Q 値学習は時間差分法の拡張である。特に Q 値の特徴として、モデルなしの意思決定に十分に適用できること、および報酬のフィードバックから直接的に学習が可能であることが上げられる。図 4 - 2 の Q 値学習で用いた変数の一覧を図 4 - 3 に示す。これらの変数と共に、強化学習の説明を行う。

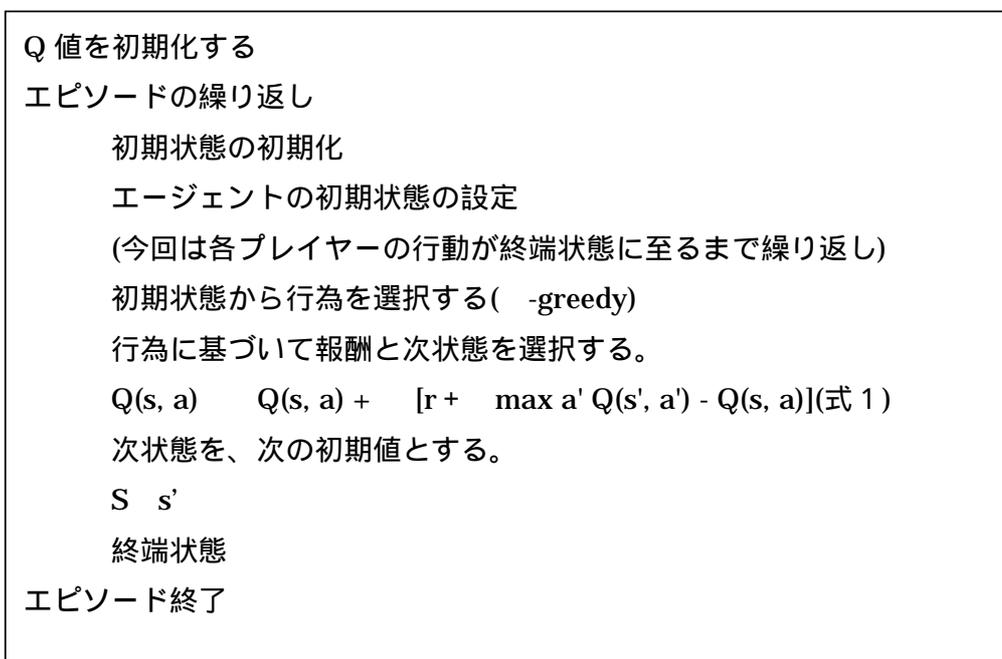


図 4-3 Q 値学習

$s=s(t)$	現状態
$a=a(t)$	現状態から選択された行為
$s'=s(t+1)$	次状態
$a'=a(t+1)$	次状態から選択された行為
	discount-rate parameter 割引率
$r$	$a$ を選択した場合に得られる報酬
	step-size parameter 学習率

図 4-4 強化学習の変数

$\max_a Q(s', a)$  は行為選択であり、 $s'$  の状態から最大の  $Q$  値となる行為を選んだときの価値関数である。次に、 $\gamma$  は割引率と呼ばれ、次のステップで最適と思われる好意を選択したときに得られると見込まれる評価の見積もりから、一段階引いた値にするために必要な定数である。1 に近いほど未来の値を推定し、0 に近いほど至近の数値を推定する。

また、 $\alpha$  は学習率であり、過去の価値を現在の価値にどの程度反映させるかを示す変数である。

$\epsilon$  は確率分岐を行う変数であり、行為決定法のひとつである  $\epsilon$ -greedy に用いる。 $0 < \epsilon <= 1$  の間に  $\epsilon$  は設定され、 $\epsilon$  の確率だけランダムな行為を選択する。逆に  $1 - \epsilon$  の確率で Greedy な(最大の価値が得られる)行為を選択する。

今回、いくつかある強化学習の枠組みの中で  $Q$  値学習を利用したのは行為と価値の組み合わせにより価値関数を更新するため、サプライチェーンのように変化が激しい問題には、より細かい評価が可能であると想定したためである。

#### (a) 状態の決定方法

本研究のシミュレーションでは、週の初めにビールゲームに関する実際の情報(注文や製品の在庫情報などの実状態)をコピーする。これにより、今週の情報および製品の移動といった状態と、ある行為を選択した後の次状態の比較に基づき、SCM におけるプレイヤー(エージェント)の状態を決定する(図 4 - 5)。

実状態( $V_t$ )を特定するために用いる情報は、サプライチェーン全体のグローバル情報として、

- ・ 他プレイヤーの在庫(3)  
を設定し、各プレイヤーのローカル情報として
- ・ 各プレイヤーの在庫
- ・ 受注残務
- ・ 顧客からの注文
- ・ 上流からの製品供給(配送分)
- ・ 顧客への発送分
- ・ 在庫から受注残務を差し引いた値

とする。在庫と受注残務の差を設定したのは、業者が判断材料として在庫と顧客からの予約から単純に導くことが往々にして存在するため、意思決定に有効であると想定した。これら9つの変数について、今週分と行為によって生成される次週の情報が見る値の差を計算し、その正負により状態を決める。なぜなら、単なる離散事象では、環境の状態は非常に幅広くなる可能性があるためである。現在の状態とその時点から1週間前の状態の比較(増減)によって現在の状態とし、次回の状態と今回の状態の比較を次の状態と設定する。このように定式化することで、2の9乗=512通りの状態を設定することが可能となった。

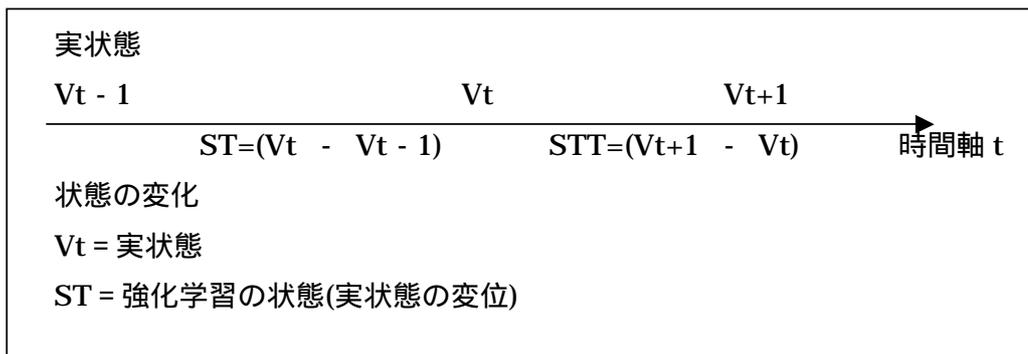


図 4-5 状態の決定

## (b) 行為選択について

強化学習で示される方法の中で、今回は特に行為選択方法のひとつである  $\epsilon$ -greedy を用いる。変数として  $(0 < \epsilon < 1)$  を設定し、閾値として  $\epsilon$  を超えた場合、次状態から選ばれた Q 値の中で最大値を選択する行為を求め、閾値を超えない場合は ランダムに行為を選択する。エージェントの行為選択のそれぞれについて、環境の状態について試行錯誤を繰り返す行為を情報獲得行為(Exploration)、既に学習された情報に基づいて価値の最大となる行為を選択するものを報酬獲得行為(Exploitation)と呼ぶ。今回は  $\epsilon$  の値をエピソードごとに変更することで、情報獲得行為中心から報酬獲得行為中心へと移行する(図 4 - 6)。

$$\epsilon = 0.1 - (0.11 / (EP - RT)) * (C - RT)$$

EP = 全エピソード数

C = 現在のエピソード

RT = 完全にランダムな回数 = EP \* (2 / 5)

図 4-6 行為選択の条件

まず、完全にランダムな行為選択を、全試行回数の 2/5 だけ続ける。

続いて、その後のエピソードでは全試行回数に到達する直前に完全に報酬獲得行為に移行するように  $\epsilon$  の値を設定する。

これにより、たとえば全試行回数が 100 回とするならば、常にランダムな回数は 40 回である。また  $\epsilon$  が 0 に到達する回数は 95 回である。

## (c) 報酬の設定

強化学習において、行為に対する報酬は在庫を表すものと受注残を表すものを同時に考慮しなくてはならないことを表している。そこで、ビールゲームにおける動的な変化を示す値から、一意に行為の妥当性を表す値を構築する必要がある。

本研究のシミュレーションでは、在庫および受注残はコスト(金)がひとつの指標

であると想定した。特に各プレイヤーの各時間単位における総コストから、その週末までの延べ総コストを導出した。さらに、サプライチェーンの指標のひとつとして、顧客満足度を設定する。通常のサプライチェーンであれば、顧客満足度をアンケートなどのデータを用いて把握している。シミュレーションでは、顧客満足度を顧客に到達した製品数と同等と想定する。

報酬の設定をもっとも単純な全体価値として、

報酬 = サプライチェーン全体の延べ総コスト / 顧客に到達した製品数

とした。これは、コストは負の数であり、顧客満足度は正の数であるからである。つまり、顧客満足度を最大化し、コストを最低にすることは、報酬の最大化につながるためである。この設定により一意に報酬が決まる。

#### (d) 変数の設定

強化学習の枠組みによって、シミュレーションを実行するために、その他の変数を設定する。

エージェントは各業者を担当するプレイヤーを割り当てる。プレイヤーにとって行為とは、発注量を決めることである。シミュレーションでは、発注量を注文量の何倍にするかによって選択する。発注量を決定する式として、

$$\text{発注量} = X * (\text{そのプレイヤーの注文量}) + Y * 4$$

と設定した。ここで X と Y の組み合わせは  $X = \{0, 1, 2\}$  とし、 $Y = \{0, 1\}$  としたため 6 通りである。これは小売業を例に挙げると、最低 0 ケース(発注せず)とし、最大で 20 ケース(=  $8 * 2 + 4 * 1$ )を発注する。これはヒューリスティック関数の場合と同等であるので、比較が可能である。

その他の変数として、 $\alpha = 0.99$ 、 $\beta = 0.01$  とする。

## 4.4 まとめ

本章では、第三章で求めた概念モデルに基づいて、ビールゲームという問題の様々な性質を含んだシミュレーションモデルとその枠組みを構築した。

特に、各プレイヤーの意思決定支援システムとして、ヒューリスティック関数と強化学習の枠組みを適用した。ヒューリスティック関数は、各プレイヤーに関する情報だけを扱うために、協働作業を行わないSCMを表現する。一方、協働作業を行う場合、サプライチェーン全体に関する情報として、他のプレイヤーの在庫情報が判明する。強化学習の枠組みで各プレイヤーの意思決定を支援することにより、協働作業を行うSCMを実現できるものとみなす。

このように、本研究におけるシミュレーションの位置付けは、知覚される現実世界としてヒューリスティック関数、概念モデルに基づく変革が可能であるという変換を実現する強化学習の枠組みを比較する場として提供されたことがわかる。

第五章では、本章で構築したシミュレーションモデル、およびその枠組みを用いたシミュレーションを実行する。ビールゲームにおける各プレイヤーの発注量を、ヒューリスティック関数および強化学習によって意思決定させたシミュレーション結果から、ビールゲームにとって有効で実行可能な改善案を提示する。

## 第五章

# シミュレーションの結果および考察

本章では、シミュレーションの実行結果、およびその考察を行う。

はじめにヒューリスティック関数を用いたシミュレーション結果を示す。次に、強化学習で用いる報酬の設定をSSMに基づいて設定し、実行結果を示す。最後に、結果からSSMにおける分析を行うことで、SCMにとっての可動性・効率性・有効性を示す。これにより、知覚される現実問題を表すデータをヒューリスティック関数が導出し、強化学習によるデータにより有効な改善案を提案する。

### 5.1 ヒューリスティック関数の場合

すべてのプレイヤーに、ヒューリスティック関数を適用したSCMのシミュレーションの結果を示す。情報遅れのない場合、各プレイヤーが100週間で生じさせるコストは、表5.1より{R(小売業), W(卸業), D(配送業), F(工場)} = {\$1,124, \$2,090, \$2,744, \$3,302}であり、小計\$9,160となる。さらに延べ総コストで示すと、\$544,606であった。

表 5-1 サプライチェーンコスト結果(情報遅れなし)

	顧客に 到達した 製品数	小売業	卸業	配送	工場	総コスト	延べ 総コスト
Heuristic	772	-1124	-2090	-2744	-3202	-9160	-544606
強化学習	772	-738	-2322	-1452	-1264	-5776	-231812

表 5-2 サプライチェーンコスト結果(情報遅れあり)

	顧客に 到達した 製品数	小売業	卸業	配送	工場	総コスト	延べ 総コスト
Heuristic	768	-1712	-3342	-4808	-5702	-15564	-805754
強化学習	764	-1112	-948	-1930	-2636	-6626	-280316

一方情報遅れのあった場合、表 5 - 2 より各プレイヤーから生じたコストは{R(小売業), W(卸業), D(配送業), F(工場)} = { \$1,712, \$3,342, \$4,808, \$5,702}であり、小計\$15,564 となる。のべ総コストは\$805,754 であった。

ここで、100 週間分の各プレイヤーが生じさせるコストと共に、100 週間のすべての延べ総コストを示した理由としては、在庫および受注残の影響は各プレイヤーだけでなく、他のプレイヤーにも影響を及ぼすためである。そこで、最終結果のみならず、各週のコストを総計することにより、サプライチェーン全体の評価に用いた。実際に、100 週間後の総コストが示す値は、後に示す強化学習の値と比較しても半分以上減ることはない。しかし、延べ総コストのヒューリスティック関数と強化学習の場合を比較するとわかるように、50%以上サプライチェーン全体のコストを削減できた。

次に、図 5 - 1 および図 5 - 2 は、ヒューリスティック関数の場合の在庫と受注残の履歴を示している。縦軸はケース数であり、正なら在庫量、負なら受注残に当たる。この図から、情報遅れのある場合は情報遅れのない場合に比べて、明らかに上流からの製品到達時間が遅れる。そのため顧客の需要は上流行程に行くにつれて増幅するため、情報遅れのない場合に比べて、余分な発注を行っている。たった一時間単位(ここでは一週間)の違いではあるが、コストの増幅は情報遅れのない場合に比べて 1.5 倍に達している。この点から、情報遅れの存在しないように、情報システムを構築することが重要である。

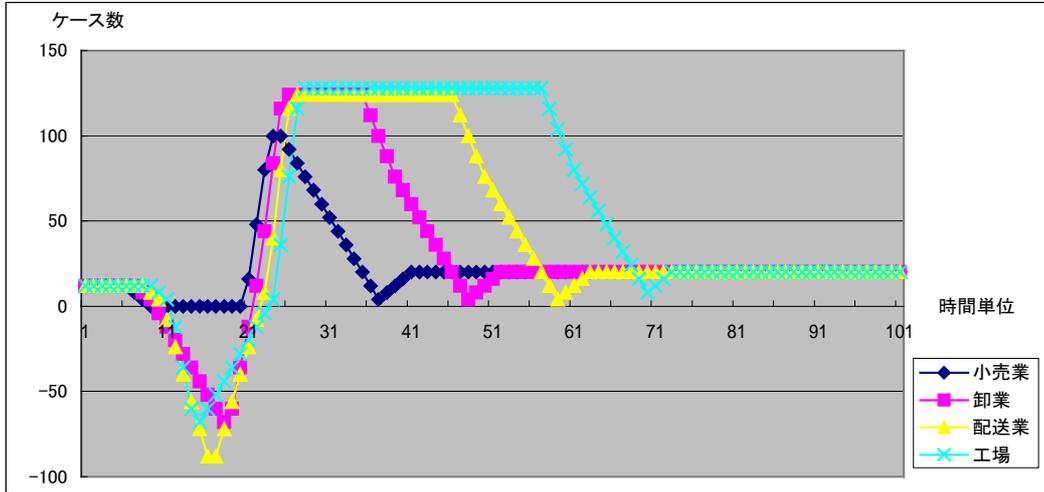


図 5-1 ヒューリスティック関数の 在庫 - 受注残 履歴(情報遅れなし)

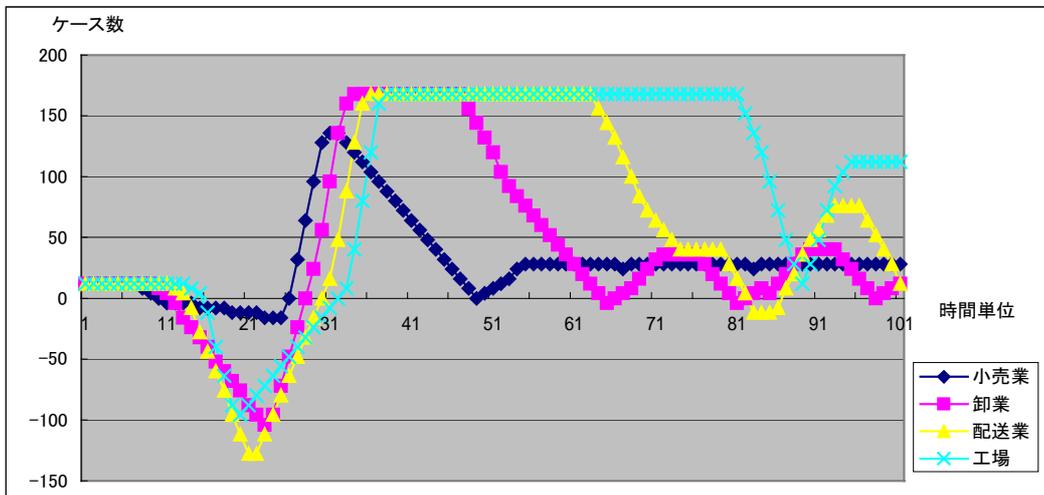


図 5-2 ヒューリスティック関数の 在庫 - 受注残 履歴(情報遅れあり)

## 5.2 強化学習の適用結果

シミュレーションの1時間単位を1週とし、100週を1エピソードとして30万エピソード試行した。最終的には完全に報酬獲得行為になるため常に最低のコストかつ最大の顧客満足度になる。

まず図5-3～図5-6は、情報遅れのない設定のデータである。一方、図5-7～図5-10は情報遅れのある場合のデータである。図5-3と図5-7は、在庫と受注残の履歴を示す。在庫から受注残をひいた数値で表す。また図5-4および図5-8では、在庫と受注残履歴の全プレイヤー平均を示す。図5-5および図5-9では、各プレイヤーの各週ごとのコストを示す。最後に図5-6および図5-10では、ヒューリスティック関数と強化学習の場合の、コストの比較を行っている。

### 5.2.1 情報遅れのない場合

情報遅れのない設定のデータから考察する。表5-1が示すように各プレイヤーのコストは、{R(小売業), W(卸業), D(配送業), F(工場)} = { \$738, \$2,322, \$1,452, \$1,264 } であり、小計\$5,776となる。この値は同じ条件のヒューリスティック関数の場合と比べてみると、0.631倍にコストが圧縮されていることを示している。さらに延べ総コストで示すと、\$231,812であった。この数値は、ヒューリスティック関数のデータと比べて0.426倍にコストが圧縮されている点から、最終結果より途中経過で効率的な経営行為(発注量の決定)がなされたことが言える。

図5-3の在庫と受注残の履歴から、前半50週までは各プレイヤーは在庫および受注残務をなるべく抑えようと努めている。一方50週を越えるころから、工場・配送はなるべく在庫を蓄積し、卸業は受注残を増やしている。しかし変動の拡大傾向は続いているにもかかわらず、小売業はほとんど目立った動きをせず、小口の受発注を行っていた。このことを、次の図5-4の在庫-受注残の平均値と合わせて見ると、サプライチェーン全体の在庫と受注残の履歴はバランスをとっており、需

給バランスは収束していることがわかる。図 5 - 5 に示した各週のプレイヤーのコストをみると、前半の 50 週と後半の 50 週の違いが明確になる。この図では、前半では各プレイヤーはなるべくコストを出さないように意思決定を行っているが、後半では在庫を蓄えた配送・工場以上に受注残が発生している卸業が大きなコストを負担している。サプライチェーン全体として卸はボトルネックになっている可能性もある。このように卸業に過大な影響が伴う理由としては、下流である小売業の変化の早さと、上流である配送および工場から送られる製品の流通にもまた時間遅れが生じるためである。このことから、卸のような中間プロセスにこそ改善点があることが見受けられる。図 5 - 6 では、ヒューリスティック関数の場合と強化学習の場合の総コスト比較を行っている。このグラフから、強化学習において卸業がボトルネックであることが明確になった。しかし、その他のプレイヤーの総コストは、大幅に向上させることができた。

## 5.2.2 情報遅れのある場合

次に、情報遅れがある場合のデータについて考察する。表 5 - 2 が示すように各プレイヤーのコストは、 $\{R(\text{小売業}), W(\text{卸業}), D(\text{配送業}), F(\text{工場})\} = \{\$1,112, \$948, \$1,930, \$2,636\}$  であり、小計  $\$6,626$  となる。この値は同じ条件のヒューリスティック関数の場合と比べてみると、0.426 倍にコストが圧縮されていることを示している。さらに延べ総コストで示すと、 $\$280,316$  であった。この数値は 0.348 倍にコストが圧縮されている。このように、コストの圧縮だけで見ると、情報遅れがない場合より大きい。情報遅れによる影響を、在庫情報の共有によるサプライチェーンの協働作業が実現できたことにより緩和したと想定する。

図 5 - 7 に示した在庫 - 受注残履歴の様相は、情報遅れのために在庫および受注残務のピークは、情報遅れのない場合に比べて数週間後にずれているが、工場および配送業は、在庫を蓄え受注残務をなるべく持たないようにしている。また、情報遅れがない場合と比べると、上流行程が在庫と受注残について大きな変動を起こしているのに対し、小売業および下流し業はあまり変動が大きくなっていない。とくに、卸業は情報遅れがない場合に比べても、また小売業と比べても変動が小さくなっている。ここで、図 5 - 8 から、在庫 受注残の平均は、情報遅れがない場合と

同じく、在庫と受注残の間で振幅を繰り返している。もちろん、情報遅れがある設定の場合、受注残より在庫のほうが多い傾向が見られる。このことは、強化学習はサプライチェーンに情報遅れがあると知覚した場合、在庫をサプライチェーン全体で、情報遅れのない場合に比べて多めにする傾向を学習したことがわかる。

強化学習によって、在庫 受注残の平均を表すグラフが、情報遅れのある場合とない場合で同様の傾向を示したことは、サプライチェーン内の協働作業により在庫と受注残を振幅の収束として表せる可能性があることである。つまり、現実のSCMにおいても、情報の共有化によって、一時的な変動を緩やかだが一定の振幅として表すことが有効ではないかと想定する。

図5-9が示す各週のコストでは、やはり工場および配送業のコストは大きな負担となっている。しかも、情報遅れなしの設定では見られなかった、小売業のコスト増が起きている。

最後に図5-10では、情報遅れのある場合の、各プレイヤーの延べコストを示している。このグラフは、すべてのプレイヤーのコストはヒューリスティック関数の場合よりよくなっている。

結果についてまとめると、

- ・ 情報遅れのあるなしに関わらず、強化学習はコストを大幅に削減する
- ・ サプライチェーン全体の在庫 - 受注残平均における振幅の発生
- ・ 情報遅れなしにおける、卸業のコスト増
- ・ 情報遅れありにおける、工場および配送のコスト増

が挙げられる。

次節では、本節で示したシミュレーション結果から、ビールゲームに対する改善案を提示する。

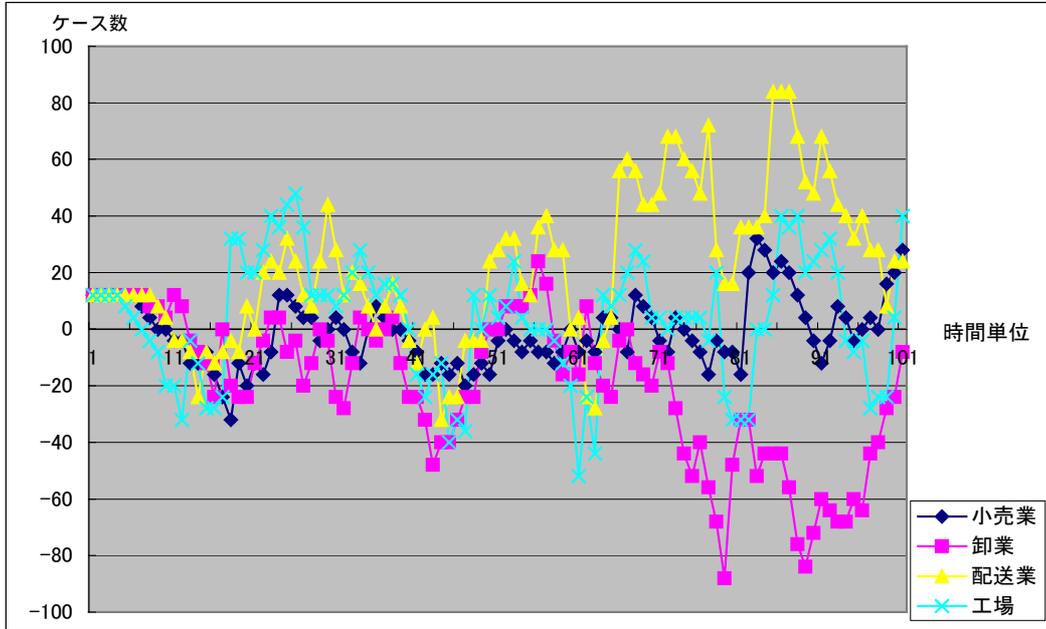


図 5-3 強化学習を適用した場合の `在庫 - 受注残` 履歴(情報遅れなし)

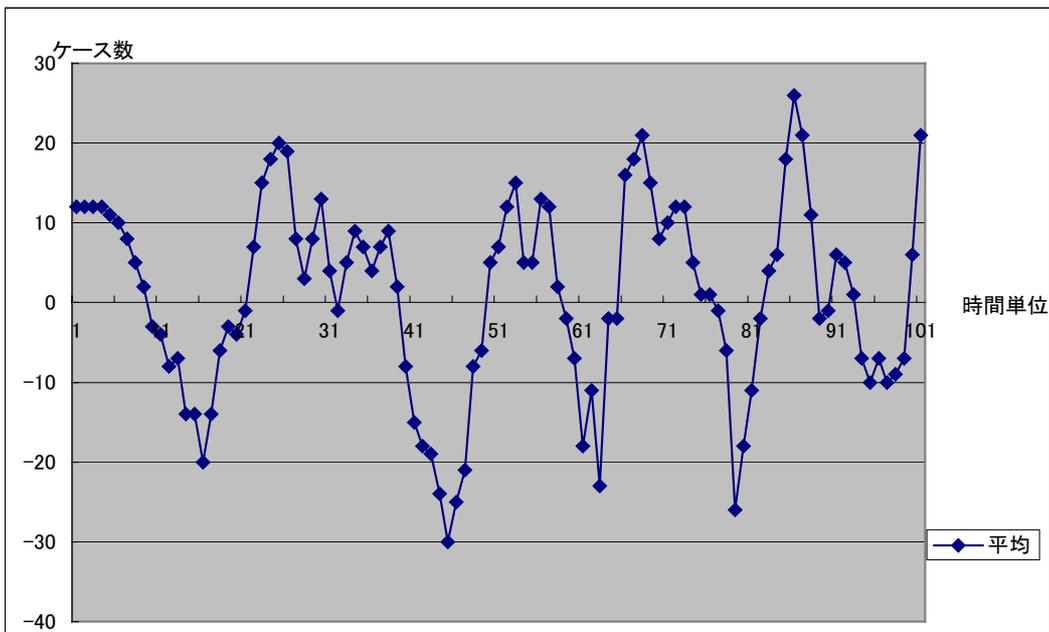


図 5-4 `在庫 - 受注残` のサプライチェーン平均(情報遅れなし)

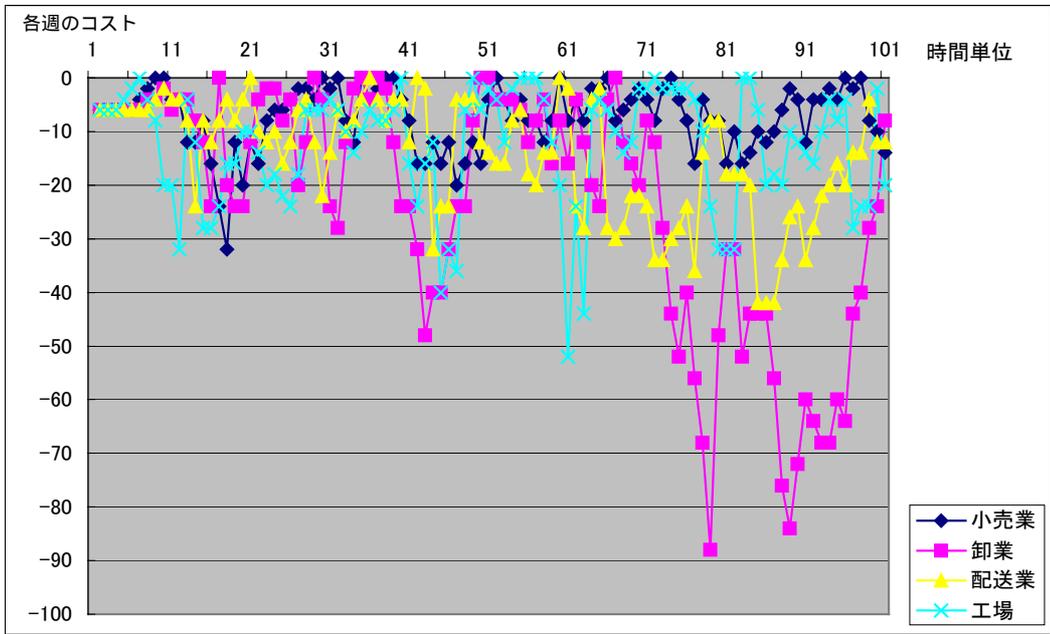


図 5-5 各週のコスト小計(情報遅れなし)

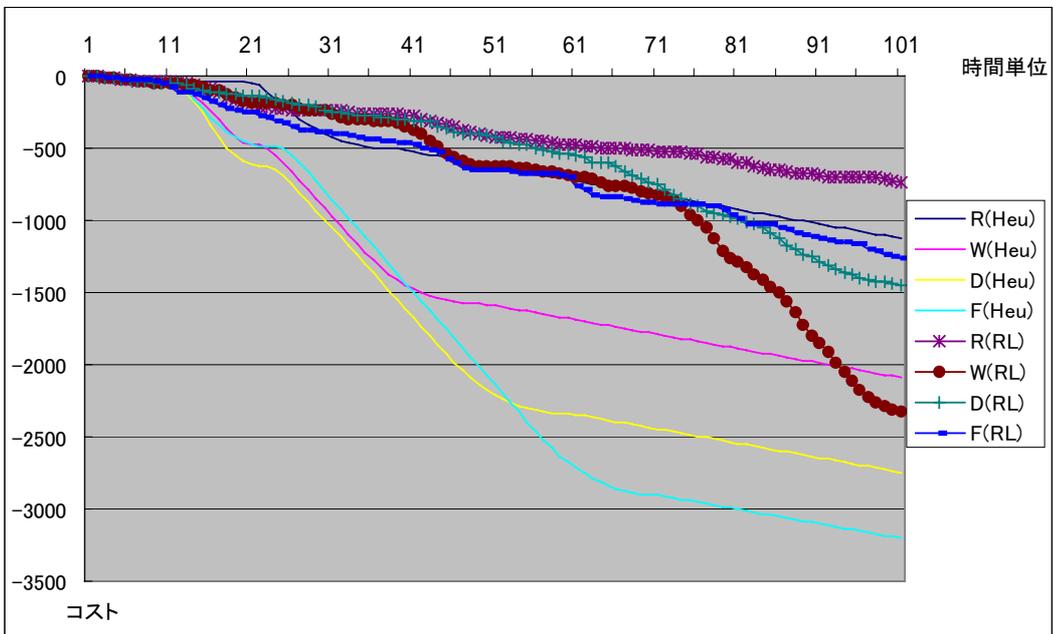


図 5-6 コストの履歴(RL:強化学習,Heu:ヒューリスティック, 情報遅れなし)

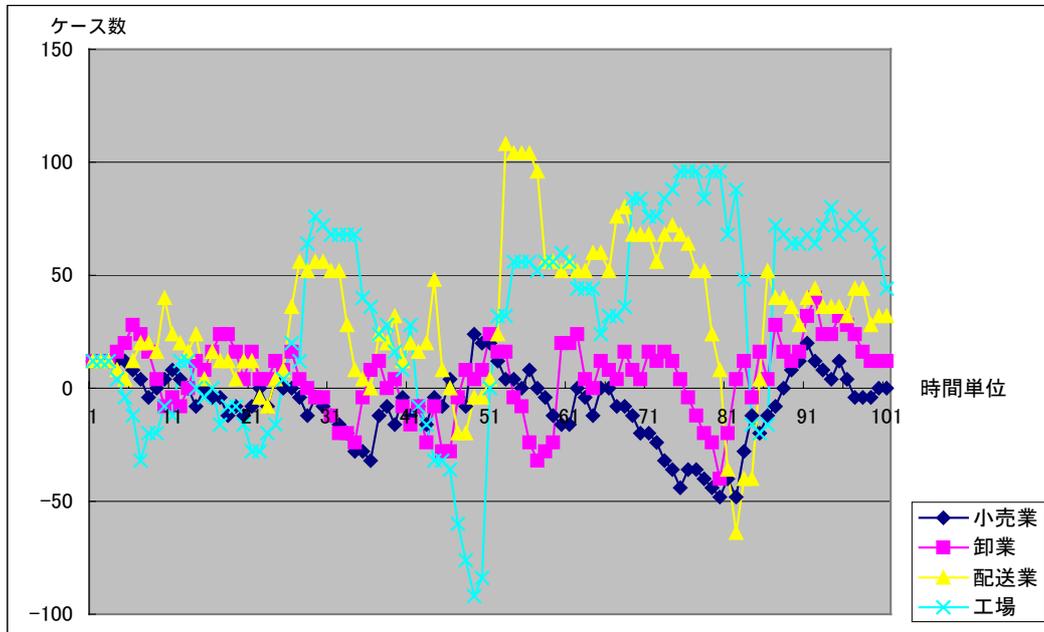


図 5-7 強化学習を適用した場合の「在庫 - 受注残」履歴(情報遅れあり)

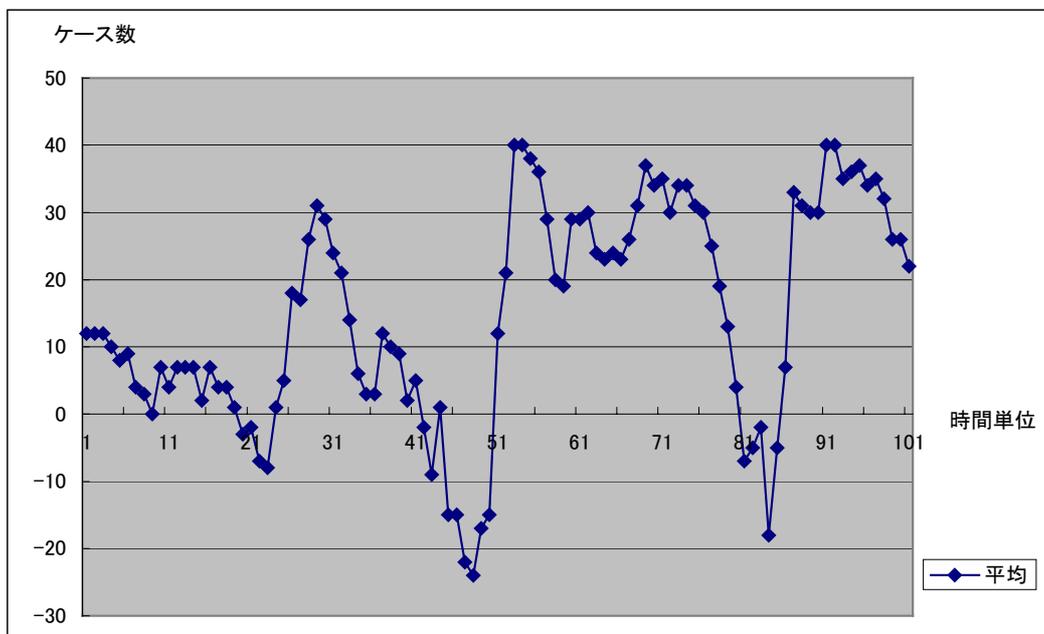


図 5-8 「在庫 - 受注残」サプライチェーン平均(情報遅れあり)

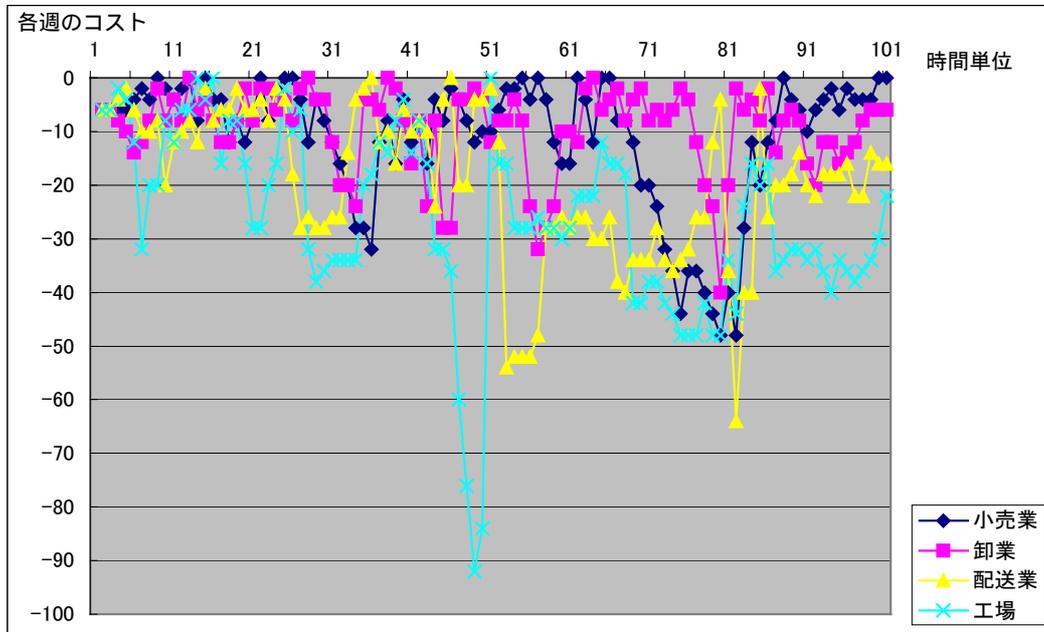


図 5-9 各週のコスト小計(情報遅れあり)

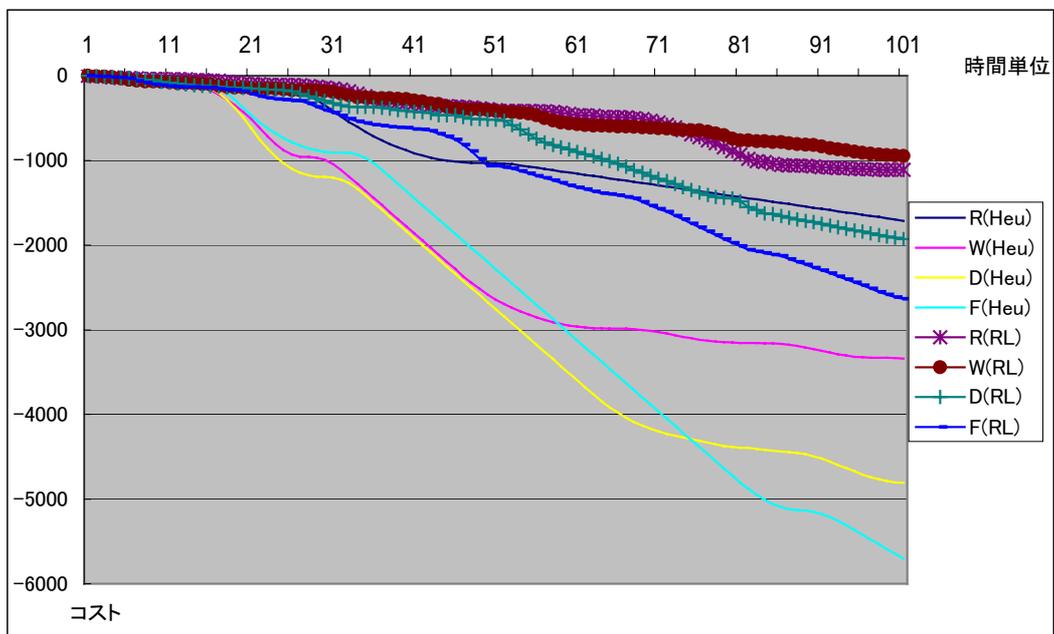


図 5-10 コストの履歴(RL:強化学習,Heu:ヒューリスティック, 情報遅れあり)

## 5.3 シミュレーション結果の考察

この節では、本研究の強化学習に対する総合的な評価、およびビールゲームに対する改善案についてまとめる。そのためにSSMにおける、論理的かつ文化的分析を行う。

次の節では、各プレイヤーに関する改善策の提言を示す。

### 5.3.1 プレイヤーごとの改善案

情報遅れの条件に関わらず、各プレイヤーの意思決定の過程から、各プレイヤーに対してどのように改善すべきかを示す。

まず、図5-3および図5-7から、顧客に最も近い小売業に求められる行為とは、受注残および在庫を抑えることであり、小口の受発注を目指すことである。情報遅れがある場合、ない場合にかかわらず、顧客の需要に合わせた発注行動を取らないといけない。もちろんサプライチェーン全体の情報にあわせて、発注量を適正に決める必要がある。

卸業では、上流からの製品流通時間および下流からの情報伝達時間の影響を受ける。特に小売業の発注は顧客の動きに連動しているため、振幅の波長は短い。そのような卸業に必要な選択としては、受注残も在庫も持たないことである。卸業は、上流および下流に対する受け皿となるように在庫をもてる存在ではあるが、通常は上流と下流の流れをせき止めないように行為選択を行うことが望ましい。

シミュレーション結果から、配送業と工場の振る舞いは非常に密接であった。そこで、配送と工場を一まとまりのグループとしてみなす。シミュレーションの傾向として、なるべく適正な在庫を持つようにして、受注残務を持たないようにすることを示している。図5-3のデータでは卸業が変動の拡大の影響を受けて、在庫の量は工場より多くなっているが、情報遅れの発生している図5-7において配送業および工場の在庫のほうが、卸業より在庫が増幅している。このことは、下流に製品を発送しても時間がかかるため、在庫を確保しておくことで、下流の変動を上流

で処理することを提案しているといえる。

### 5.3.2 ビールゲームの論理的分析

一般に、強化学習として問題を表すことの困難さは、環境・強化信号・学習率や割引率などの変数をどのように決定するか、という部分にある。SCMの問題は、物理現象と異なり恣意性が多い。つまり、ある状態に何らかの行為を行ったからといって、に適切な評価を行えるとは限らない。特に一般的なサプライチェーンは、在庫コストや納期、顧客満足度といったいくつかの指標が存在する。またそれぞれの指標はトレードオフの関係にある。つまり一方の価値をあげることは他方の価値を下げることに繋がることがありうる。

本研究では、サプライチェーンを評価する指標として、顧客に到達する製品数を顧客満足度とし、在庫と受注残務から生じるコストからコスト対顧客満足度の最大化を目指してきた。最終的には今回のシミュレーションでは顧客満足度である製品の到達数についてはほぼ可能であった。コストのみを比較した場合でも、情報遅れがないときで総コストのほぼ40%削減(延べ総コストに至っては50%以上削減)できた。さらに強化学習による意思決定は、情報遅れに伴う変動の拡大傾向を抑え、50%以上の高いコスト削減が達成された。

ビールゲームは、変動の拡大という性質を示しているため、SCMの一面を表している。そこで、SCMを基本定義に示したように変換できたかどうかは、ビールゲームの経済性・有効性によって示せる。第三章において定めたように、経済性はコストであり、有効性は顧客満足度で表せる。コストの最小化、および顧客満足度の最大化はビールゲームのシミュレーションで示した。これによりビールゲームの経済性、有効性は向上した。

概念モデルでは、協働作業にすることで情報を共有することが期待でき、それにより、受注残および在庫の削減が導けるようにSCMを変革できるというプロセスを構築した。実際に、シミュレーションによって比較すると、プレイヤー間の情報共有を実現した場合、強化学習のシミュレーションが、現実世界をあらゆるヒューリスティック関数を適用した場合の経済性および有効性を向上させている。

強化学習を用いた離散事象シミュレーションが、ビールゲームの経済性および有

効性などのサプライチェーンのパフォーマンスを向上させるための、実行可能で有効な改善案を導出できた。

今回、強化学習を適用したことで、情報遅れのない場合のボトルネックは卸業であり、情報遅れのあるほうでは、工場および配送業がボトルネックとなる可能性がわかった。情報遅れがあるほど上流行程に影響が大きい。そこで、ビールゲームに対する改善案として、

- ・ 情報遅れはなるべく避ける。情報がスムーズに流れるように、情報システムを改善すべきである。
- ・ 情報遅れがある場合、上流行程はあまり下流の影響をリアルタイムに受けない。
- ・ 情報遅れがない場合でも、卸業の在庫量に注意してサプライチェーン全体の発注量を調整すべきである。

ということが言える。

本研究において、強化学習の枠組みは、SCMにおいて顧客満足度の最大化およびコストの最小化を行った。これにより、強化学習の枠組みはSCMの意思決定支援システムとして有効である。さらに、SSMによって報酬の設定、モデルの構築のための情報が概念モデルによって実際の世界と比較できたので、ビールゲームにおけるSCMの変換は、この節で示した改善案により議論可能な意見を導けた。

### 5.3.3 ビールゲームの文化的分析

本研究では、サプライチェーンの所有者を明確には設定していない。ビールゲームは、ロールプレイングシミュレーションゲームであるため、現実のSCMの一面しか表せないためである。今後の課題として、実際の企業への適用による検証が必要である。

## 第六章

### 終わりに

#### 6.1 総括

本研究の主題は、SCM を強化学習の枠組みでシミュレーションモデルに表すことであった。また、様々な問題の組み合わせさせたシステムである SCM を、シミュレーションモデルとして実装するためのシステム方法論について明確にし、これらの方法論がもつシステミックな世界観の重要性を示すことにあった。特に、サプライチェーンに対する TSI の適用は、適切なシステム方法論として SSM を提案することを示した。またシミュレーションが、方法論における現実世界との比較および有効な改善案の提案に用いることができることを示した。これにより、TSI や SSM といった方法論は、SCM の有効性・効率性・妥当性を検討する学習システムとして人間活動システムを考察する場合適切であることが示された。

本研究で示した、方法論による問題の明確化、シミュレーションによる系統的に望ましく文化的に実行可能な変革の提示、自己システムの中で組織学習を繰り返すことによる環境の変化に対するアジルな(機敏)対応、これらの方法論は企業組織の思考速度を市場に合わせることを要求している。世界的に企業文化の変革スピードが上昇している昨今、組織や企業は学習システムを導入することにより、経営に対して広い視野を持つべきである。学習システムを企業グループが確立し、かつシミュレーションなどのスモールワールドに対する評価法を見出すことが、顧客の変化、コストの変化、供給の変化、技術の変化に応じた社会を形成できる。

## 6.2 今後の課題

強化学習の枠組みによりビールゲームの問題を扱ってきたが、強化学習を適用するにあたっていくつかの問題点が現れた。

まず、強化学習の枠組みで設定した報酬は、概念モデルの意味から導いたものである。この報酬の有効性は、経済性から適切であったが、偶然によって導かれたものかもしれない。また、キャッシュフローや顧客の動向・受発注量・材料については、本研究では無制限に設定している。このような設定は、実際の提案を行う際には妥当性を欠く理由となる。今後、ビールゲームのようなSCMの離散事象シミュレーションでは、より有効性を高めるために、条件を厳しくするなど行うべきである。

また、SCMにおけるサプライチェーンの勝利とは、サプライチェーンに携わる企業群がすべて勝たないといけない、つまり負け組みを作ってはいけないことにある。強化学習を用いることで全体コストは下げることができたが、さらにサプライチェーンの弱い部分を補強できるように、強化学習の枠組みを変更する必要がある。

また、強化学習の設定についても

- ・ 報酬の設定
- ・ サプライチェーン有効性の指標
- ・ 適正な在庫
- ・ 適正な供給

の構築について、より実際に即した深い考察が必要である。

このような妥当性、有効性、効率性を検証するために、さらにSSMの学習サイクルにのっとり、考察を繰り返す必要がある。

一方、SSMを用いるときには、ある程度の経験が必要となる。今回の強化学習によるシミュレーションにおいては、経験を事前の試行錯誤によって補っている。このような制約を取り払うひとつの方法としては、文献[9]で指摘されている、最良の経営スタイルのひとつであるミドル・アップダウン・マネジメントがある。知識創造プロセスは、ミドル・マネジャを知識マネジメントの中心として認識し、企業内

情報のタテとヨコの流れが交差する場所に位置づけている。これらのマネジャはしばしばチームやタスクフォースのリーダーである。彼らの役割は業務改革のための提言およびそれらの実装である。現状としては、そこにはシステミックアプローチと呼べるようなものは認識できないが、成果物は具体的なものとなっている。このように、彼らおよび彼らの活動こそ、人間活動システムである想定できる。すなわち、ミドル・アップダウン・マネジメントの概念をシステム思考やSSMと調和させることは意義深い。これが達成されることにより、適用における豊富な経験の要求を多少は排除できるものと考えられる。今後の研究に期待したい。

また、シミュレーションにおける強化学習の課題と共に、SSMについてもより深いシステム思考によって、取り組まなくてはならない。また、今後は実践的な方法論であることを示すために、文化的論理的分析について取り組む必要がある。

## 謝辞

本研究を実行するに当たり、尽力頂きました吉田武稔助教授に感謝いたします。経営学および強化学習の基本について示していただいた、田中雄介氏および内藤浩行氏に感謝いたします。SSMの別の見方であるWSRを紹介していただいたGu Jifa教授に感謝いたします。共に学びさまざまな機会に援助いただいた吉田研究室一同様に感謝いたします。

## 参考文献

- [1] ダイヤモンド・ハーバードビジネス編集部, “サプライチェーン理論と実践”, ダイヤモンド社, 1998.
- [2] Supply Chain Council inc, “SCOR Overview of the SCOR Model v3.0”, sep 1998.(この文献は以下のサイトでダウンロードできます;  
[http://www.supply-chain.org/html/scor\\_overview.cfm](http://www.supply-chain.org/html/scor_overview.cfm))
- [3] Robert L. Flood, and Michael C. Jackson, “Creative Problem Solving: Total Systems Intervention”, John Wiley & Sons Ltd., 1991.
- [4] ピーター・チェックランド, “新しいシステム方法論”, オーム社, 1985.
- [5] ピーター・チェックランド. ジム・スクールズ(妹尾監訳), “ソフト・システムズ方法論”, 有斐閣, 1994.
- [6] Richard S. Sutton, and Andrew G. Barto, “Reinforcement Learning”, The MIT Press, 1998.
- [7] Sterman J.D., “Modeling managerial behavior: misperceptions of feedback in a dynamic decision making experiment”, Management Science, Vol. 35, No. 3, March 1989.
- [8] Sterman J.D., “Deterministic chaos in an experimental economic system”, Journal of Economic Behavior and Organization, Vol. 1, No. 28 North - Holland, Dec 1989.
- [9] 野中,竹内(梅本訳), “知識創造企業”, 東洋経済新報社, 1996.
- [10] センゲ(守部訳), “最強組織の法則”, 徳間書店, 1995.
- [11] ボイエット,ボイエット(金井監訳), “経営革命大全”, 第3章, 日本経済新聞社, 1998.

- [12] 吉田武稔, 田野勇二, 田中雄介, "サプライチェーンマネジメントのモデル化に関する考察", 計測自動制御学会システム情報部門シンポジウム, 高知工科大学, 11月11, 12日, 1999.