

Title	動的信念論理に基づく信念変更のAIへの応用
Author(s)	上島, 駿平
Citation	
Issue Date	2015-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/12657
Rights	
Description	Supervisor: 東条 敏, 情報科学研究科, 修士

修 士 論 文

動的信念論理に基づく信念変更の AI への応用

北陸先端科学技術大学院大学
情報科学研究科情報科学専攻

上島 駿平

2015 年 3 月

修 士 論 文

動的信念論理に基づく信念変更の AI への応用

指導教員 東条 敏 教授

審査委員主査 東条 敏 教授
審査委員 Nguyen Minh Le 准教授
審査委員 飯田 弘之 教授

北陸先端科学技術大学院大学
情報科学研究科情報科学専攻

1310031 上島 駿平

提出年月: 2015 年 2 月

概要

近年人工知能という分野が非常に発展してきている。人工知能とは人間の知能を持った機械であり、その研究分野は広く、遺伝アルゴリズム、音声認識、画像認識、感性処理、機械学習、自然言語処理、論理といった様々な観点から研究が行われており、日常生活において幅広く実用化されている。また人間の脳を解析し、精密な思考を行える人工知能が生み出すことができれば、遠くない未来において人工知能が更に賢い人工知能を造るという予想がされている。その人工知能を作り出すためには、人間がものごとを考えていくプロセスを形式化することが必要となってくる。論理学の分野において、人間の知識を扱う論理が盛んであり、その一つとして動的信念論理がある。

従来、数理論理学における動的信念論理は、論理学者が人間が新しい情報を獲得するときに、今まで持っていた知識・信念をどのように変更するか、といった問題を人工知能の分野への応用とは独立に研究していた。信念とは、特定の人物が信じている内容を指すものであり、本人の勘違いや思い込みを含めている。また、知識とは、客観的に見たときの事実である。信念は認識している世界においてのみ真という性質上、変更を可能にしている。信念が変更される方法として拡張、修正、コントラクションがある。拡張とはある信念に対して情報が与えられたとき、信念集合はその信念と情報との和集合となる。また、修正とは新しい情報を受け入れる代わりに信念の中の命題を一部削除し無矛盾性を保つものであり、コントラクションとは、新しい情報を加えないで信念の中の命題を削除し一貫性を保つものである。信念変更はこれらの方法を用いることで、無矛盾性を維持しつつ信念の動的な変化を観察することを目的としている。

そこで本研究の目的は、今まで基礎的数学的関心に閉じている動的信念論理を現実の人工知能の問題に応用することである。そのために動的信念論理を用いて知識・信念の更新を行うコンピュータシステムを実装することにより、推理小説や論理パズルにおける可能世界の削減過程をシミュレートし、検証を行うことが必要である。人間が推論を行う過程をコンピュータ上で実現しようとするとき、推論を行うシステム部分と、推論に必要な情報を保持しておく知識ベースが必要となる。知識ベースには、現在正しいと信じている情報とそれが何故正しいのかという根拠が含まれていて、新しい情報が加えられると、それに応じて正しいと信じている情報を生成しなおす仕組みが一般的である。知識・信念の更新は情報の累積ではない。それらが更新していく過程をシミュレートするために、可能性のある選択肢の削除という観点で捉えることにより、情報の獲得によって変化する信念状態を表すことを目的とする。

論理パズルの例として sum and product というものがある。これは、エージェント間の会話によって情報が更新されていき、到達可能性が次第に削除されていくものである。一見して情報がなさそうな告知を行うが、その実は大量の到達可能性を消去している。この問題を、信念が会話によって更新されていく過程をモデル化することができる。そして推理小説では、物語の進行に伴い推論材料となる情報を読者が得ることにより真相に辿り

着くことができる。読者の推論は、文章中の因果関係や時間的順序関係といった事柄に対して論理的構造を構築するということである。推理小説における犯人推論を、文章中の有意性を持つ命題を基に、読者の知識・信念が更新されていくプロセスと捉えることが可能であり、inquisitive semantics の発想を用いことで、選択肢の削除を行う。

Inquisitive semantics の発想を用いた動的信念論理である、Inquisitive dynamic epistemic logic がある。従来、可能世界を削ることに情報量があり、有意味であるとされてきた。しかしこの論理では可能世界を削ることのない告知にも意味を持たせている。 $p \vee \neg p$ のような告知をしたとしても従来の論理においては告知はトートロジーであり、すべての可能世界で成り立つので可能世界の数は変わることはなく、この論理式には情報量はない。しかし少なくとも p について関心があるということがいえる。その関心を抱いている部分を issue を用いて形式化することが可能となる。

さらに Inquisitive dynamic epistemic logic に対し、新しく古典的否定演算子 $\sim$ を導入することで、今まで行えなかった推論が可能になる。この論理を IDEL $\sim$ としてを提案する。この論理では可能世界の集合をグループとして捉ることにより、そのグループの中のどこかの世界で成り立っている命題に対して意味づけを行える。また各グループで成り立っている命題に対して推論を行っていることにより、可能世界の数が爆発することを防いでいる。そして推理小説の犯人推論において複数犯の可能性がある場合を形式化することに適して

これらの例を用いて、知識・信念の更新をするコンピュータシステムを実装することにより、可能世界の削減過程をシミュレートし検証する。

目次

第1章	はじめに	1
1.1	研究の背景	1
1.2	研究の目的	1
1.3	本稿の構成	2
第2章	人間の信念を扱ったモデル	3
2.1	信念を扱ったパズル	3
2.1.1	sum and product	3
2.1.2	誕生日の謎	4
2.2	推理小説における人間の信念変化	5
第3章	動的認識論理	10
3.1	基礎的な論理	10
3.1.1	命題論理	10
3.1.2	述語論理	13
3.2	様相論理	17
3.2.1	構文論	17
3.2.2	クリプキ意味論	18
3.2.3	論理式と到達可能性関係の	20
3.2.4	様相論理のヒルベルト流公理系	22
3.2.5	自然演繹	23
3.3	公開告知論理	25
3.3.1	構文論	25
3.3.2	意味論	26
3.3.3	クリプキモデルの例	28
3.3.4	誕生日の謎	28
第4章	Inquisitive semantics	33
4.1	Information states と issue	33
4.1.1	Information states	33
4.1.2	Issues	34
4.2	Inquisitive Dynamic Epistemic Logic	35

4.2.1	構文論	35
4.2.2	意味論	37
4.2.3	告知演算子によるモデルの書き換え	43
第 5 章	IDEL[~] の提案とその実装と解析	45
5.1	実装	45
5.1.1	システムの構成	45
5.1.2	誕生日の謎をシステムを用いての検証	49
5.2	推理小説の解析	51
5.2.1	IDEL を用いた可能性の削減	51
5.2.2	IDEL [~]	54
5.3	IDEL [~] の汎用性	61
第 6 章	終わりに	65
6.1	考察	65
6.2	今後の課題	65

目 次

2.1	人物相関図	8
3.1	クリプキモデルの例	19
3.2	設定したクリプキモデル	28
3.3	変更後のクリプキモデル	28
3.4	到達可能性関係	30
3.5	初期状態	31
3.6	一つ目の告知により変化した状態	32
3.7	二つ目の告知により変化した状態	32
3.8	三つ目の告知により変化した状態	32
4.1	Information states	34
4.2	Issues	34
4.3	issue の初期状態	42
4.4	告知による Σ_a の変化	44
5.1	モデルの設定	47
5.2	Frame	47
5.3	初期状態	48
5.4	告知の真偽値	48
5.5	モデルの更新	48
5.6	告知と変化した Frame	50
5.7	modifyFrame1	50
5.8	modifyFrame2	50
5.9	modifyFrame3	50
5.10	各可能世界とその issue	52
5.11	issue の要素	52
5.12	ann1 後の Σ_r	53
5.13	proposition $\llbracket \sim \varphi \rrbracket$	56
5.14	初期状態の p-issue	57
5.15	ann1 後の p-issue	57
5.16	ann2 後の p-issue	58

5.17 ann2 後の可能世界と p-issue	60
5.18 ann2 後の Σ_r	60
5.19 ann3 後の可能世界と p-issue	61
5.20 ann3 後の Σ_r	61
5.21 ann4 後の可能世界と p-issue	62
5.22 ann4 後の Σ_r	62
5.23 初期状態	63
5.24 告知 $m \wedge k$ 後の issue	63
5.25 告知 $\sim \neg m \wedge \sim \neg k$ 後の p-issue	64
5.26 告知 $\neg h_b$ 後の p-issue	64

表 目 次

2.1	犯行可能性	9
3.1	φ, ψ の複合命題の真理値表	11
3.2	論理式と到達可能性関係	20

第1章 はじめに

1.1 研究の背景

従来、数理論理学における動的信念論理 [1] は、論理学者が、人間が新しい情報を獲得するときに、今まで持っていた知識・信念をどのように変更するか、といった問題を人工知能の分野への応用とは独立に研究していた。信念とは、特定の人物が信じている内容を指すものであり、本人の勘違いや思い込みを含めている。また、知識とは、客観的に見たときの事実である。信念は認識している世界においてのみ真という性質上、変更を可能にしている。信念が変更される方法として拡張、修正、コントラクションがある [2]。拡張とはある信念に対して情報が与えられたとき、信念集合は信念と情報との和集合となる。また、修正とは新しい情報を受け入れる代わりに信念の中の命題を一部削除し無矛盾性を保つものであり、コントラクションとは、新しい情報を加えないで信念の中の命題を削除し一貫性を保つものである。信念変更はこれらの方法を用いることで、無矛盾性を維持しつつ信念の動的な変化を観察することを目的としている。

1.2 研究の目的

本研究の目的は、動的信念論理を用いて知識・信念の更新を行うコンピュータシステムを実装することにより、推理小説や論理パズルにおける可能世界の削減過程をシミュレートし、検証を行うことである。人間が推論を行う過程をコンピュータ上で実現しようとするとき、推論を行うシステム部分と、推論に必要な情報を保持しておく知識ベースが必要となる。知識ベースには、現在正しいと信じている情報とそれが何故正しいのかという根拠が含まれていて、新しい情報が加えられると、それに応じて正しいと信じている情報を生成しなおす仕組みが一般的である。知識・信念の更新は情報の累積ではない。それらが更新していく過程をシミュレートするために、可能性のある選択肢の削除という観点で捉えることにより、情報の獲得によって変化する信念状態を表すことを目的とする。

今まで基礎的数学的関心に閉じている動的信念論理であるが、現実の人工知能の問題に応用することが必要である。例えば推理小説では、物語の進行に伴い推論材料となる情報を読者が得ることにより真相に辿り着くことができる。読者の推論は、文章中の因果関係や時間的順序関係といった事柄に対して論理的構造を構築するということである。推理小説における犯人推論を、文章中の有意性を持つ命題を基に、読者の知識・信念が更新されていくプロセスと捉えることが可能であり、Ciardelli, Groenendijk, Roelofsen らによって

提案された inquisitive semantics [3] の発想を用いることで、選択肢の削除を行う。また、論理パズルの例として “sum and product” [4] というものがある。これは、エージェント間の会話によって情報が更新されていき、到達可能性が次第に削除されていくものである。一見して情報がなさそうな告知を行うが、その実は大量の到達可能性を消去している。この問題を、信念が会話によって更新されていく過程をモデル化することができる。これらの例を用いて、知識・信念の更新をするコンピュータシステムを実装することにより、可能世界の削減過程をシミュレートし検証する。

1.3 本稿の構成

本論文では、第2章において人間の信念がどのように扱われているのかについて例を用いて記述する。一つ目の例として sum and product とその簡易版としての “誕生日の謎”。二つ目の例として、推理小説における犯人推論である。

第3章、第4章では、第2章で持ち上げた例に対して表現できる論理について説明していく。第3章では動的信念論理の一種である公開告知論理について、その元となる命題論理、述語論理、様相論理について学んだことをまとめていき、公開告知論理を用いて “誕生日の問題” について解法を与える。第4章では、inquisitive semantics について説明をしていき、その応用例を与える。

5章では、新しく提案した論理 IDEL[~] について説明し、その実装について述べ、推理小説について解析、検証を行い、評価実験とそれに対する考察を行い、6章にて本論文のまとめと今後の展望について示していく。

第2章 人間の信念を扱ったモデル

本章では，人間の信念がどのように扱われているのかについてパズル，小説の2つの面から示していく．

2.1 信念を扱ったパズル

人間の信念を扱った例として sum and product [4] とその簡易版としての“誕生日の謎”について書いていく．

2.1.1 sum and product

Sum and product とは 1969 年にオランダの数学誌 *Nieuw Archief voor Wiskunde* の中で H. Freudenthal によって発表したのが初出である．この問題は動的信念論理での研究の応用として用いられ，Hans van Ditmarsch [5] らにより，形式化が行われてきた．以下が日本語に翻訳したものになる．

A は S と P に対しての発言：「2つの数字 x と y を次の条件を満たすように選んできた．その条件は $1 < x < y$ であり， $x + y \leq 100$ である．そこで S には二つの数字を足した答え $s = x + y$ を， P には二つの数字を掛け合わせた答え $p = x \cdot y$ を教える．教えた内容は互いに秘密である．」彼は言った通りに二人に教え，そして次の会話が起った：

1. P の発言：「数字の組み合わせは分かりません．」
2. S ：「そうだと知っていました．」
3. P ：「いま，数字の組み合わせが分かりました．」
4. S ：「私も分かりました．」

x と y の組み合わせ (x, y) はなんだったでしょう．

この P と S の発言は一見何の情報も持っていないかのようにみえるが，その一つ一つの発言により大量の可能性が絞られていく．現に P と S は相手の発言を聞いてすぐ答えが分かっている．

では実際に発言の意味を考えて推論してみよう．まず一つ目の推論の， P の発言であるが， P にとって分かる組み合わせとはなにか考える必要がある． P には二つの数字のかけ合わせた数を知っているのだから，素数同士の組み合わせであれば瞬時にその組 (x, y) が分かってしまう．また，大きすぎる組み合わせになると $x + y \leq 100$ という条件により分かってしまう組もある．例えば P が教えられた数字が 2499 であれば $(49, 51)$ と分かってしまう．

次に二つ目の推論． S は P が分からないということを知っていたので，足してできた数 s を $x + y$ の形に分解したとき， x と y は P が分からないような数の組でなければならない． S が教えられた数字 s が 15 であれば， $x + y$ に分解したときに $(2, 13)$ という組み合わせができてしまい，これでは P が分かるはずなので S はそのことを知っていたとは発言できないのである．しかし s が 11 であれば $x + y$ に分解してもその二つの数は素数同士にはならないので今のところ正しいと言える．またこの推論により可能性は大幅に減るのだがここまでのステップだけでは私たちは答えに辿り着くことはできない．

次に三つ目の推論． P は S の発言を受けて組み合わせが分かった言っている．二つ目までの推論によって二つの数を足した数が絞れていることは分かっている．掛けた数 p を $x \cdot y$ に分解したとき，その二つの数を足した数 s というのは上で推論した二つの数になるので，18 であれば $(2, 9)$ に分解することができ，その二つの数を足した数 s は 11 になりこの組だと P は二つの数が分かる．また p が 30 であれば分解したときに $(5, 6)$ ， $(2, 15)$ の組み合わせがあり，その組の s は 11 と 17 になる．二つ目の推論により絞られた数 s にこの 2 数は満たしているので，この p では P は数字の組が分かったとはいえない．

そして最後の推論となる． S も会話により正解の組み合わせが分かる． $(2, 9)$ であれば三つ目までの推論を満たす．しかし S が一意に決められるわけではない． $(2, 9)$ を足してできた数 11 を分解すると他にも $(3, 8)$ や $(4, 7)$ があるのだが，この組み合わせであっても三つ目までの推論を満たしてしまうのである．このようにして推論していけば，答えは足して 17，掛けて 52 になる $(4, 13)$ しか残らないことが分かる．以上のような思考の過程を辿り人間は回答を得ることができる．

2.1.2 誕生日の謎

このパズルは端的にいうと sum and product を簡単化したパズルであり，考え方は sum and product とほとんど同じであるが，探索空間が sum and product ほど大きいものではない．本質は sum and product と同じ推論の仕方であるので可能性の数が少ない sum and product として“誕生日の謎”についてこれから説明していく．

A さんと B さんは C さんの誕生日にプレゼントをあげたいのですが，二人とも C さんの誕生日の日付がわかりません．ただ C さんの誕生日は下の 10 組の中のどれかだけわかります．

- 3月4日

- 3月5日
- 3月8日
- 6月4日
- 6月7日
- 9月1日
- 9月5日
- 12月1日
- 12月2日
- 12月8日

CちゃんはAくんに「何月」誕生日か教えました、そしてBくんに「何日」誕生日か教えました。

- Cちゃん「さあ、私の誕生日はいつですか？」
1. Aくん「俺は答えがわからないから、Bくんもわからないはずです」
 2. Bくん「俺もわからなかったけど、今のAくんの言葉でわかりました」
 3. Aくん「あっ、じゃ俺もわかりました」

ではCちゃんの誕生日は10組の中のどの日でしょうか？

この問題の解答へ至る推論の方法を記述していく。

まずBくんがCちゃんの誕生日が分かるであろう日付は2日であれば12月2日、7日であれば6月7日と一意に決まってしまうので、その二つの日付はありえない。そしてAくんはその日付が思い浮かぶ日付というのはいない、12月と6月ではないことが分かる。Bくんはこの会話によって分かったので、いまだ区別がつかない3月5日と9月5日でないことが分かる。最後にAくんも分かったので、答えは9月1日になる。

この「誕生日の謎」の推論の構造はsum and productと同様であり、段階的発言によりその発言が嘘になってしまう組み合わせである日付を除外していくことに意味がある。最初はこの問題を解いていく人間からすれば、3月4日から12月8日までの10個の可能性があるのだが、AくんとBくんの会話によりだんだんとCちゃんの誕生日である可能性のある日付というものが変化していく。そして最後に残った可能性というのが答えになるのである。

2.2 推理小説における人間の信念変化

推理小説においても先の二つの例であげたように可能性が次第に消されていって推論していく解き方を消去法として一般的に知られている。コナンドイルの生み出した世界一有名であろう名探偵であるシャーロックホームズも自らの推理法を

「不可能な物をすべて除外してしまえば、あとに残ったものが、たとえいかに不合理に見えても、それこそ真実に違いない」(When you have eliminated the impossible, whatever remains, however improbable, must be the truth.)

と『白面の兵士』[6]の中で述べている。

しかし逆に日本の推理小説作家・坂口安吾は『不連続殺人事件』[7]の附記としての選後論争の中で以下のように消去法について斥ける発言をしている。

「いわば探偵小説のトリックとは、消去法を相手にして、それによる限り必ず失敗するようにつくられたものである。消去法によるとまっさきに犯人でなくなってしまうような完全なアリバイをもつ人物が、実は犯人であるという、そこにトリックがあり、探偵小説の吟味があるのである。」

上のシャーロックホームズの発言は小説中のキャラクターの発言であり、坂口安吾の発言は作者自らの発言であり、その違いはあるが、二つとも推理小説の消去法についての記述である。

二つの文は一見矛盾しているかのように見えるかもしれないが、どちらも真であるといえる。シャーロックホームズによる発言はすべての可能性について考え、それを少しずつ削っていった残る可能性がほぼ一つに絞られてしまう。これは推理小説においては大事な要素である。犯人が探偵によって明示されるが、その他にも可能性が残っている人物というのが残ってはいないものになる。推理小説において明確にその犯人にしか行えないのであれば、すべてのほかの可能性について文中で否定できる要素を描く必要があるのである。推理小説が本当にフェアプレイをしているのであれば、すべての読者は理詰めにより犯人を言い当てることができるのである。

また坂口安吾の発言も正しく、真っ先に疑わしい人物というのは一旦可能性から消えたように思えるが、その読者に可能性を削らせる過程にトリックがあるのである。トリックが一切なく、その可能性が残った人物というのが犯人であればあまりにも面白みがない。作中でトリックを用いることで、初めて意外な犯人が現れるのである。シャーロックホームズの発言とは、その消えたかのようにみえる可能性についても十分に裏を取り、本当に可能性がないのかと疑ってかかり、そしてトリックを見破った先に見えてくる可能性についても考慮しており、またそのトリックについても十分に論理的である必要性もある。よって、二文というのは推理小説の本来持っている性質について語っているものであるといえる。

では実際に、坂口安吾の『不連続殺人事件』を例にとって、その可能性が消えたようにみえるトリックについて検証していこうと思う。

不連続殺人事件

『不連続殺人事件』は坂口安吾の初めて書いた推理小説であり、1947年に大地書房発行の雑誌『日本小説』に連載されていた。掲載時には作者による、読者への挑戦として真

犯人あての懸賞金がかけられていた。本書についてのあらすじは角川書店の単行本の裏表紙を引用して説明したい。

戦後間もないある夏，詩人・歌川一馬の招待で，山奥の豪邸に集まったさまざまな男女。作家，詩人，画家，劇作家，女優など，いずれ劣らぬ変人・奇人ぞろい。邸内に以上な愛と憎しみが交差するうちに，世にも恐るべき，8つの殺人が生まれた！不連続殺人の裏に秘められた悪魔の意図は何か？鬼才案後が読者に挑んだ不滅のトリック！多くのミステリ作家が絶賛する，日本推理小説史に輝く傑作。第2回探偵作家クラブ賞受賞作。

あらすじの通り，殺されるのは8人であり，登場人物は40人余り，容疑者の数だけでも30人近くいる。図2.1がこの小説の登場人物の相関図である。見て分かれるとおり，歌川多門を中心にいろいろな職業の男女が色恋により，色濃く関係し合っている。

この推理小説を解析するためには，すべての事件においてのアリバイや証言について解析，また人物関係についてくまなく探り出し，すべての可能性を考えた上で，一筋の筋書きを整えて犯人に迫る必要があるが，あくまでここでは論理的な解き方の一例を示すことができる。とする。

よって小説の肝となるトリックを単純化することで小説に対する論理的な回答を与えることにする。単純化したものが以下の通りになる。

容疑者はA, B, C, D, E

ある村で殺人事件が3つ起こる

一つ目の殺人のとき，A, Bは村の外に出ており，目撃者にも確認が取れていて確固としたアリバイがある。DとEは共に一緒にいたと証言。Cのみ部屋におりアリバイがない。

二つ目の殺人のとき，怪しいとされるC, D, Eはそれぞれ自分の部屋にいて，Bがそれを歩きながら廊下で監視していた。Bは3人とも部屋を出ることはなかったと証言。C, D, EもまたBが廊下を離れることはなかったと証言。Aのみアリバイがない。

三つ目の殺人のとき，A, C, Dは一緒にいたと互いに証言した。BとEにはアリバイがなかった。

犯人は一人もしくは，その共犯が一人だけいる可能性がある。さあ犯人は誰。

この問題を解くために必要なのは証言の正確さである。まずすべての容疑者の発言が真だとして，最初の事件では犯行可能なのはCだけ，二つ目の事件ではAだけ，三つ目の事件ではBとEになる。表2.1は，殺人可能である容疑者には○を，殺人不可能な人には×をもって表している。

このとき犯人が一人だとしたら，第一の事件によってA, B, D, Eにはアリバイがあり，Cのみが捜査線上に浮上する。しかし，第二の殺人において，Cにはアリバイがある

表 2.1 犯行可能性

	A	B	C	D	E
第一の殺人	×	×	○	×	×
第二の殺人	○	×	×	×	×
第三の殺人	×	○	×	×	○

ため容疑者から外され、容疑者がいなくなる．ここで考えなければいけないのは問題として与えられた条件である共犯についてである．いま犯人に共犯がいるとすると、最初の事件では C が犯行を行い、また二つ目の事件では A が犯行を行うことができる．だが、第三の事件において A,B ともにアリバイがあり、犯行は不可能になってしまう．さらにこの図のどの二人の組み合わせをみても、犯行を行える組み合わせはなくなってしまう．

そこで各容疑者の証言について考える．第一の事件において D と E は共に一緒にいたと証言している．このとき D と E は共犯であればこの証言について偽ることによりアリバイを作ることができる．が D と E の組み合わせであると、第二の殺人において D,E のアリバイを B は証言している．共犯は一人だけいると文中に記されているため B の証言は真であり、これも可能性としては残らない．

次に C が第一の殺人を行った場合を考える．C と A が共犯である場合、第一、第二の事件はクリアされるが第三の事件はどうか．A と C は共に証言しあっているが、ここにまたしても第三者である B が介入しているため、このアリバイは成り立ち A と C という組み合わせもありえない．

では B と C はどうか．第一、第三の事件では犯行を行える．第二では B は C, D, E によってそのアリバイを保障されているので犯行は行えないであろう．しかし C についてはどうか．C, D, E の 3 人によってアリバイを保障されている B であるが、C について嘘をつくことは可能なのである．ここにトリックがあり、第三者によってアリバイを保障された人物でも嘘の証言をつくことにより共犯のアリバイを保障しているのである．よって第二の殺人で B の証言は嘘であったとすると、B と C の共犯である可能性が残る．

また C, D や C, E といった組み合わせも考えられるが、第二の事件で共に B に証言されているため、アリバイがあり犯行は行えない．可能性を削っていった最後に残った組み合わせは B と C だけであるので、この問題の解は B と C の共犯になる．

第3章 動的認識論理

人間の信念を表す論理体系として、様相論理の一種として認識論理 (Epistemic Logic) がある。様相論理は、「必然性」や「可能性」について論じられるロジックとして、アリストテレスの時代から研究が始められ、哲学者や神学者の長年の積み重ねにより現代に受け継がれている。この様相論理には様々な解釈を与えることができ、時制論理や知識の論理である認識論理もその一つである。認識論理によって人間が「信じている」や「知っている」といった命題を様相演算子として捉えることができ、これを形式化することができる。また、その知識・信念の変化を情報の更新と捉えてその過程をシミュレートすることができるのが動的認識論理 (Dynamic Epistemic Logic) [1] である。

本章では命題論理、述語論理、様相論理を踏まえた上で、動的認識論理の一種である公開告知論理について書いていく。

3.1 基礎的な論理

3.1.1 命題論理

構文論

論理式を構成するための記号を示していく。

命題変数 命題を記号化、形式化したものとして命題変数 (propositional variable) を定義。これは文の最小単位である。

$p, q, r, \dots$

論理式 命題は論理結合子を用いて組み合わせることで複合命題を作り出すことができる。

- $\neg$: 「… でない」 否定 (negation)
- $\wedge$: 「かつ」 論理積 (conjunction)
- $\vee$: 「または」 論理和 (disjunction)
- $\rightarrow$: 「… ならば …」 含意 (implication)

BNF (Backus Naur Form) 記法で構文を定義すると以下ようになる.

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \varphi \wedge \varphi \mid \varphi \rightarrow \varphi$$

意味論

真理値表 論理結合子について日本語で「かつ」などを意味していることが分かったが, それは結合子の意味を与えたことにはならない. この意味を理解するのに便利なのが真理値表である. 本稿では真を 1, 偽を 0 で表すことにする.

表 3.1 φ, ψ の複合命題の真理値表

φ	ψ	$\neg\varphi$	$\varphi \wedge \psi$	$\varphi \vee \psi$	$\varphi \rightarrow \psi$
1	1	0	1	1	1
1	0	0	0	1	0
0	1	1	0	1	1
0	0	1	0	0	1

「かつ」という接続詞で二つの命題 φ と ψ を繋いだ後にできる命題 $\varphi \vee \psi$ は φ, ψ が共に真であるときにのみ真であることを示している. $\neg\varphi$ は φ が偽のとき真になり, $\varphi \wedge \psi$ は φ と ψ のどちらか片方でも真であれば真といえる. $\varphi \rightarrow \psi$ は前件が真でかつ後件が偽のとき以外が真となる.

真理値表を用いることで, 原始命題の真偽値が定まれば複合命題の真偽値も同時に決まることが分かる.

付値 命題変数の与え方を扱うために, 論理式全体の集合から真偽値 $\{1, 0\}$ に対する写像を付値とする.

以下では「 P がなりたつための必要十分条件 (iff: if and only if) は Q がなりたつことである」ということを表すために, メタ記号 $\iff$ を用いて,

$$P \iff Q$$

と書くことにする.

表 3.1 の論理式を付値を用いて表すと,

1. $v(\neg\varphi) = 1 \iff v(\varphi) = 0$
2. $v(\varphi \wedge \psi) = 1 \iff v(\varphi) = 1 \text{ かつ } v(\psi) = 1$
3. $v(\varphi \vee \psi) = 1 \iff v(\varphi) = 1 \text{ または } v(\psi) = 1$

$$4. \quad v(\varphi \rightarrow \psi) = 1 \iff v(\varphi) = 1 \quad \text{または} \quad v(\psi) = 1$$

論理式 φ において、全ての付値 v に対し $v(\varphi) = 1$ がなりたつとき φ はトートロジー（恒真）といふ。

証明論

自然演繹はゲンツェンが導入した形式化の一つであり、公理系を最小に留める代わりに推論規則を豊富に用意したものである。推論規則は通常数学で行われる推論に近い形であり、理解しやすい自然な体系であるといえる。

推論規則 以下に示すような、論理演算子に関する「導入規則 (Introduction)」と「除去規則 (Elimination)」と、例外的な $\perp$ に関する規則、背理法 (RAA) に関する規則がある。

$$\begin{array}{c}
 \frac{\varphi_1 \wedge \varphi_2}{\varphi_i} (\wedge E) \quad \frac{\varphi_1 \quad \varphi_2}{\varphi_1 \wedge \varphi_2} (\wedge I) \quad \frac{\varphi \vee \psi \quad \begin{array}{c} [\varphi]_1 \\ \vdots \\ \chi \end{array} \quad \begin{array}{c} [\psi]_2 \\ \vdots \\ \chi \end{array}}{\chi} (\vee E)_{1,2} \quad \frac{\varphi_i}{\varphi_1 \vee \varphi_2} (\vee I) \\
 \\
 \frac{\varphi \rightarrow \psi \quad \varphi}{\psi} (\rightarrow E) \quad \frac{\begin{array}{c} [\varphi]_1 \\ \vdots \\ \psi \end{array}}{\varphi \rightarrow \psi} (\rightarrow I)_1 \quad \frac{\neg \varphi \quad \varphi}{\perp} (\neg E) \quad \frac{\begin{array}{c} [\varphi]_1 \\ \vdots \\ \perp \end{array}}{\neg \varphi} (\neg I)_1 \\
 \\
 \frac{}{\perp} (\perp) \quad \frac{\begin{array}{c} [\neg \varphi]_1 \\ \vdots \\ \perp \end{array}}{\varphi} (RAA)_1
 \end{array}$$

ここで $[\varphi]$ のように各括弧がついている論理式は仮定であり、推論規則の適用後に仮定のリストがら消される。また $\perp$ は矛盾を表す記号である。推論規則 RAA について説明を追加する。 RAA は「 $\neg \varphi$ を仮定し、矛盾が生じるときには φ が導かれる」という意味であり、背理法 (reduction ad absurdum) の頭文字を取ったものである。また命題論理の自然演繹で式 φ が定理になるとは、仮定のリストが空になり、 φ に至る証明図が存在することである。

直観主義論理は上に挙げた推論規則のうち RAA を取り除いたものである。古典論理では直観主義論理では成り立たない排中律が成り立つのだが、その証明の際に RAA の規則が必要となる。以下に排中律の証明を与える。

$$\begin{array}{c}
\frac{[A]_1}{A \vee \neg A} (\vee I) \quad \frac{[\neg(A \vee \neg A)]_2}{\perp} (\neg E) \\
\frac{\perp}{\neg A} (\neg I)_1 \\
\frac{A \vee \neg A}{\perp} (\vee I) \quad \frac{[\neg(A \vee \neg A)]_2}{\perp} (\neg E) \\
\frac{\perp}{A \vee \neg A} (RAA)_2
\end{array}$$

例 3.1 $A \rightarrow ((A \wedge B) \vee (A \wedge \neg B))$ に至る証明図は次のように与えられる.

$$\begin{array}{c}
\frac{[A]_3 \quad [B]_1}{A \wedge B} (\wedge I) \\
\frac{(A \wedge B) \vee (A \wedge \neg B)}{[\neg((A \wedge B) \vee (A \wedge \neg B))]_2} (\vee I) \quad \frac{[\neg((A \wedge B) \vee (A \wedge \neg B))]_2}{\perp} (\neg E) \\
\frac{[A]_3}{\neg B} (\neg I)_1 \\
\frac{A \wedge \neg B}{(A \wedge B) \vee (A \wedge \neg B)} (\wedge I) \quad \frac{[\neg((A \wedge B) \vee (A \wedge \neg B))]_2}{\perp} (\neg E) \\
\frac{(A \wedge B) \vee (A \wedge \neg B)}{A \rightarrow ((A \wedge B) \vee (A \wedge \neg B))} (\vee I) \quad \frac{\perp}{(A \wedge B) \vee (A \wedge \neg B)} (RAA)_2 \\
\frac{(A \wedge B) \vee (A \wedge \neg B)}{A \rightarrow ((A \wedge B) \vee (A \wedge \neg B))} (\rightarrow I)_3
\end{array}$$

命題論理の意味論と自然演繹の間には完全性・健全性が成り立っていることが知られており、すべての式 φ に対して、

$$\varphi \text{ がトートロジー} \iff \varphi \text{ が古典命題論理の自然演繹で証明できる}$$

ということがいえる。また $\implies$ を完全性、 $\impliedby$ を健全性という。

3.1.2 述語論理

述語論理は命題論理を拡張したものであり、命題の内部構造を分析することができる論理である。命題論理では「A は B である」といった述語に対して、 p などの単純命題に置き換えることで表現していたが、述語論理ではその内部構造を明示するために、述語記号と変数記号を用いて表現する。また量化記号により「すべての」や「ある」といった概念についても記述できる。

構文論

論理結合子 $\wedge, \vee, \rightarrow, \neg$

量化記号

- $\forall$: 「すべての」 全称記号
- $\exists$: 「ある」 存在記号

量化記号は一階述語論理において対象変数のみに用いられる.

項

- 変数記号 $x, y, z, \dots, a, b, c, \dots$

述語記号

- $P(x, \dots, x_n)$: (P は述語, $x, \dots, x_n$ は引数)

BNF 記法で構文を定義すると以下ようになる.

$$\varphi ::= P(a_1, \dots, a_n) \mid \neg\varphi \mid \varphi \vee \varphi \mid \varphi \wedge \varphi \mid \varphi \rightarrow \varphi \mid \forall x P(x) \mid \exists x P(x)$$

これらの構文論を用いることで, 複雑な文章を内部構造を考慮した命題で表現することができる. 例えば, 「 x は推理小説である」を $N(x)$, 「 x は死ぬ」を $D(x)$, 「 x は y の登場人物である」を $H(x, y)$ とすると, 「全ての推理小説において, ある登場人物が死ぬ」という自然言語の文が表現でき以下ようになる.

$$\forall x. \exists y. (N(x) \rightarrow (H(x, y) \wedge D(y)))$$

意味論

述語論理の論理式に対し意味づけを与えるためには, 「すべての」や「ある」を表現しなくてはならないので, 対象変数の動く範囲である, 対象領域を定めなければいけない. その対象領域のことを D とする. また, 意味づけのために解釈が必要であり, その解釈を集合論的に表したものが付値関数 V である. これら二つを合わせて構造 (モデル) $\mathfrak{M} = (D, V)$ と表され, 次のように与えられる.

- D : 空でないものの集まり
- V : 次の性質を満たす付値関数 (述語記号 P や変数記号 (a, b) にモノで「値」を与える)
 - $V(a), V(b) \in D$ (a, b は変数記号)
 - 述語記号 P の引数の数が 1 なら, $V(P)$ は D の部分集合.

$$d \in V(P) \subseteq D$$

- P の引数の数が n なら, $V(P)$ は D^n の部分集合.

$$(d_1, \dots, d_n) \in V(P) \subseteq D^n$$

述語論理のモデル $\mathfrak{M} = (D, V)$ が与えられたとき、論理式に真・偽 (1,0) を割り当てる真理条件を次のように定める：

1. $V(P(a_1, \dots, a_n)) = 1 \iff (V(a_1), \dots, V(a_n)) \in V(P).$
2. $V(P(a_1, \dots, a_n)) = 0 \iff (V(a_1), \dots, V(a_n)) \notin V(P).$
3. $V(\neg\varphi) = 1 \iff V(\varphi) = 0.$
4. $V(\neg\varphi) = 0 \iff V(\varphi) = 1.$
5. $V(\varphi \wedge \psi) = 1 \iff V(\varphi) = 1 \text{ かつ } V(\psi) = 1.$
6. $V(\varphi \wedge \psi) = 0 \iff V(\varphi) = 0 \text{ あるいは } V(\psi) = 0.$
7. $V(\varphi \vee \psi) = 1 \iff V(\varphi) = 1 \text{ あるいは } V(\psi) = 1.$
8. $V(\varphi \vee \psi) = 0 \iff V(\varphi) = 0 \text{ かつ } V(\psi) = 0.$
9. $V(\varphi \rightarrow \psi) = 1 \iff V(\varphi) = 0 \text{ あるいは } V(\psi) = 1.$
10. $V(\varphi \rightarrow \psi) = 0 \iff V(\varphi) = 1 \text{ かつ } V(\psi) = 0.$
11. $V(\forall x\varphi(x)) = 1 \iff \text{すべて } d \in D \text{ について } V(\varphi(\underline{d})) = 1.$
12. $V(\forall x\varphi(x)) = 0 \iff \text{ある } d \in D \text{ について } V(\varphi(\underline{d})) = 0.$
13. $V(\exists x\varphi(x)) = 1 \iff \text{ある } d \in D \text{ について } V(\varphi(\underline{d})) = 1.$
14. $V(\exists x\varphi(x)) = 0 \iff \text{すべて } d \in D \text{ について } V(\varphi(\underline{d})) = 0.$

ただし $\underline{d}$ は要素 $d \in D$ の名前で $V(\underline{d}) = d$ とする。

述語論理において、命題論理のトートロジーに対応するのが妥当 (valid) な論理式であり、任意のモデル $\mathfrak{M} = (D, V)$ で真になる論理式である。

証明論

推論規則 述語論理の自然演繹の推論規則は以下の通りである.

$$\begin{array}{c}
 \frac{\varphi_1 \wedge \varphi_2}{\varphi_i} (\wedge E) \quad \frac{\varphi_1 \quad \varphi_2}{\varphi_1 \wedge \varphi_2} (\wedge I) \quad \frac{\varphi \vee \psi \quad \begin{array}{c} [\varphi]_1 \\ \vdots \\ \chi \end{array} \quad \begin{array}{c} [\psi]_2 \\ \vdots \\ \chi \end{array}}{\chi} (\vee E)_{1,2} \quad \frac{\varphi_i}{\varphi_1 \vee \varphi_2} (\vee I) \\
 \\
 \frac{\varphi \rightarrow \psi \quad \varphi}{\psi} (\rightarrow E) \quad \frac{\begin{array}{c} [\varphi]_1 \\ \vdots \\ \psi \end{array}}{\varphi \rightarrow \psi} (\rightarrow I)_1 \quad \frac{\neg \varphi \quad \varphi}{\perp} (\neg E) \quad \frac{\begin{array}{c} [\varphi]_1 \\ \vdots \\ \perp \end{array}}{\neg \varphi} (\neg I)_1 \\
 \\
 \frac{}{\perp} (\perp) \quad \frac{\begin{array}{c} [\neg \varphi]_1 \\ \vdots \\ \perp \end{array}}{\varphi} (RAA)_1 \quad \frac{\varphi(a)}{\forall x. \varphi(x)} (\forall I)^\dagger \quad \frac{\forall x. \varphi(x)}{\varphi(a)} (\forall E)
 \end{array}$$

$\dagger$: a は任意の自由変数. すべての x を a で置き換える. さらに, $\varphi(a)$ を導くために使った仮定のどれにも a は出現しない.

$$\frac{\varphi(a)}{\exists x. \varphi(x)} (\exists I)$$

ただし, $(\exists I)$ において $\varphi(x)$ は $\varphi(a)$ 中のいくつかの a の出現を x で置き換えたもの.

$$\frac{\begin{array}{c} [\varphi(a)] \\ \vdots \\ \psi \end{array}}{\psi} (\exists E)^\ddagger$$

$\ddagger$: 自由変数 a は, $\exists x. \varphi(x), \psi$, さらに ψ を導くために使った $\varphi(a)$ 以外の仮定, の全部に出現しない.

ここで $[\varphi]$ または $[\varphi(a)]$ のように各括弧がついている論理式は仮定であり, 推論規則の適用後に仮定のリストがら消される. また述語論理の自然演繹で式 φ が定理になるとは, 仮定のリストが空になり, φ に至る証明図が存在することである.

例 3.2

$\exists y. \forall x. R(x, y) \rightarrow \forall x. \exists y. R(x, y)$ に至る証明図は次のように与えられる.

$$\begin{array}{c}
\frac{[\forall x.R(x, b)]_2}{R(a, b)} (\forall E) \\
\frac{[\exists y.\forall x.R(x, y)]_1}{\exists y.R(a, y)} (\exists I) \\
\frac{\exists y.R(a, y)}{\forall x.\exists y.R(x, y)} (\exists E)_{b: \text{fresh}, 2} \\
\frac{\forall x.\exists y.R(x, y)}{\exists y.\forall x.R(x, y) \rightarrow \forall x.\exists y.R(x, y)} (\forall I)_{a: \text{fresh}} \\
\frac{}{\exists y.\forall x.R(x, y) \rightarrow \forall x.\exists y.R(x, y)} (\rightarrow I)_1
\end{array}$$

述語論理の意味論と自然演繹の間には完全性・健全性が成り立っていることが知られており、すべての式 φ に対して、

$$\varphi \text{ は妥当な論理式である } \iff \varphi \text{ が古典述語論理の自然演繹で証明できる}$$

ということがいえる。

3.2 様相論理

古典論理によって数学におけるさまざまな理論を形式化することができた。しかし、我々が推論する際というのは、場面や状況、立ち位置によって意味の捉え方というものはさまざまである。また未来によって起こる可能性はあるなど不確実なことを考えるときもある。そういった状況に応じた思考の違いや可能性について論じることができる論理が様相論理である。様相論理は古典論理の論理結合子に加え、「必然性」や「可能性」を扱える様相演算子 (modal operator) があり、状況や場面や時間など設定の仕方により解釈の仕方が変わってくる。例えば、信念を扱う認識論理では「必然性」に対して「知っている」などの意味に解釈を変えることが可能である。その状況や場面や時間を設定するモデルがクリプキ (S. Kripke) によって与えられたクリプキモデルである。

3.2.1 構文論

命題論理の命題変数、論理結合子に加え、様相演算子がある。

様相演算子

- 必然性演算子 $\Box$: $\Box\varphi$ 「必然的に φ である」
- 可能性演算子 $\Diamond$: $\Diamond\varphi$ 「 φ が可能」
 - $\Box\varphi$ が現在の状況で真になるのは、エージェント a が現在の状況から想像しうるどのような状況でも φ が真である場合であり、その場合に限る

- $\Diamond\varphi$ が現在の状況で真になるのは、エージェント a が現在の状況から想像するある状況で φ が真である場合であり、その場合に限る

BNF 記法で構文を定義すると以下のようなになる。

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \varphi \wedge \varphi \mid \varphi \rightarrow \varphi \mid \Box\varphi \mid \Diamond\varphi$$

3.2.2 クリプキ意味論

様相論理の意味論はクリプキフレームにより定められる。空でない集合 W と W 上の二項関係 R の対 (W, R) をクリプキフレームといい、 W を可能世界 (possible world) の集合、 R を到達可能関係 (accessibility relation) という。

また各命題変数に対してどの可能世界で真になるかを定めておく必要があり、これをクリプキフレーム (W, R) 上の付値とする。付値 V は各命題変数 p に対し $V(p) \subseteq W$ となる写像であり、つまり $V(p)$ は p が真となる可能世界 w の集合を意味している。この三重対 (W, R, V) をクリプキモデルという。与えられたクリプキモデル $M = (W, R, V)$ に対して可能世界 $w \in W$ と論理式の間の二項関係 $\models$ を次のように帰納的に定義する。これは真理条件 (truth condition) と呼ばれる。

1. $M, w \models p \iff w \in V(p)$
2. $M, w \models \varphi \wedge \psi \iff M, w \models \varphi$ かつ $M, w \models \psi$
3. $M, w \models \varphi \vee \psi \iff M, w \models \varphi$ または $M, w \models \psi$
4. $M, w \models \varphi \rightarrow \psi \iff M, w \not\models \varphi$ または $M, w \models \psi$
5. $M, w \models \neg\varphi \iff M, w \not\models \varphi$
6. $M, w \models \Box\varphi \iff wRu$ となる、すべての u に対し $M, u \models \varphi$
7. $M, w \models \Diamond\varphi \iff wRu$ となる、ある u に対し $M, u \models \varphi$

$M, w \models \varphi$ であるとき、「可能世界 w で論理式 φ は真である」という。また、 $M, w \not\models \varphi$ は「 $M, w \models \varphi$ でない」ことを表す。関係 $\models$ は付値 V から一意的に定まるので、 V と $\models$ は同一視して、 $\models$ のことを付値ともいう。様相論理においても、述語論理と同様に命題論理のトートロジーに対応するのが、妥当 (valid) な論理式であり、任意のクリプキモデル $M = (W, R, V)$ で真になる論理式である。

クリプキモデルはより直感的に理解しやすいように図でもって表現される。以下のクリプキモデル $M = (W, R, V)$ を考える。

- 可能世界 w の集合 $W = \{w_0, w_1, w_2\}$

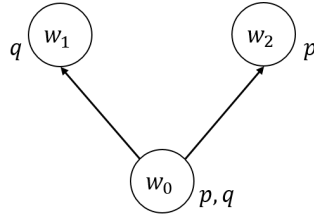


図 3.1 クリプキモデルの例

- 到達可能性関係 $R = \{(w_0, w_1), (w_0, w_2)\}$
- 付値 $V(p) = \{w_0, w_2\}$ $V(q) = \{w_0, w_1\}$

このクリプキモデルを図を用いて表すと図 3.1 となる．可能世界は w_0, w_1, w_2 の三つがあり， w_0 から w_1 ， w_0 から w_2 へと到達可能性関係がある．命題 p が真となる世界は w_0 と w_2 であり，命題 q が真となる世界は w_0 と w_1 である．このクリプキモデル上で $M, w_0 \not\models \Box(p \vee q) \rightarrow (\Box p \vee \Box q)$ が成り立つことを証明すると以下のようなになる．

例 3.3.1

$$M, w_0 \not\models \Box(p \vee q) \rightarrow (\Box p \vee \Box q) \iff \underline{M, w_0 \models \Box(p \vee q)}_{\textcircled{1}} \text{ かつ } \underline{M, w_0 \not\models \Box p \vee \Box q}_{\textcircled{2}}$$

①の証明

$$\begin{aligned} M, w_0 \models \Box(p \vee q) &\iff w_0 R u \text{ となるすべての } u \text{ に対して } M, u \models p \vee q \\ &\iff M, w_1 \models p \vee q \text{ かつ } M, w_2 \models p \vee q \\ &\iff (M, w_1 \models p \text{ または } M, w_1 \models q) \text{ かつ } (M, w_2 \models p \text{ または } M, w_2 \models q) \\ &\iff w_1 \in V(q) \text{ かつ } w_2 \in V(p) \quad \text{よって ① は成り立つ} \end{aligned}$$

②の証明

$$M, w_0 \not\models \Box p \vee \Box q \iff \underline{M, w_0 \not\models \Box p}_{\textcircled{a}} \text{ かつ } \underline{M, w_0 \not\models \Box q}_{\textcircled{b}}$$

①aの証明

$$\begin{aligned} M, w_0 \not\models \Box p &\iff w_0 R u \text{ となる，ある } u \text{ に対して } M, u \not\models p \\ &\iff u = w_1 \text{ とすると } M, w_1 \not\models p \\ &\iff w_1 \notin V(p) \text{ より ①a は成り立つ} \end{aligned}$$

⑥の証明

$$\begin{aligned}
M, w_0 \not\models \Box q &\iff w_0 R u \text{ となる, ある } u \text{ に対して } M, u \not\models q \\
&\iff u = w_2 \text{ とすると } M, w_2 \not\models q \\
&\iff w_2 \notin V(q) \text{ より ⑥ は成り立ち, ② が成り立つ}
\end{aligned}$$

よって $M, w_0 \not\models \Box(p \vee q) \rightarrow (\Box p \vee \Box q)$ が証明された.

様相論理における妥当な論理式とは任意のフレーム (W, R) , 任意のモデル (W, R, V) で真になる論理式である.

3.2.3 論理式と到達可能性関係の

ここでは代表的な論理式を満たす到達可能性関係について表で示す.

表 3.2 論理式と到達可能性関係

論理式	到達可能性関係
T $\Box\varphi \rightarrow \varphi$	反射的 (reflexive) 全ての w で wRw
4 $\Box\varphi \rightarrow \Box\Box\varphi$	推移的 (transitive) wRw' かつ $w'Rw''$ ならば wRw''
D $\Box\varphi \rightarrow \Diamond\varphi$	継続的 (serial) 全ての w に w' が存在して wRw'
B $\varphi \rightarrow \Box\Diamond\varphi$	対称的 (symmetric) wRw' ならば $w'Rw$
5 $\Diamond\varphi \rightarrow \Box\Diamond\varphi$	ユークリッド的 (Euclidean) wRw' かつ wRw'' ならば $w'Rw''$

上記の論理式はそれぞれに強弱な関係があり, **T**であれば少なくとも自分自身に行き先があるので**D**であるといえる. またユークリッド性と反射性を同時に仮定すれば, 対称性と推移性を導く. 反射性・対称性・推移性の三つの条件を満たしているものを同値関係 (equivalence relation) という.

これらのフレームの構造によってその演算子の特性が規定され, 様相論理で導入されている様相演算子も同様にこのフレームの構造によって特性が大きく変化する. 以下に代表的な論理式を挙げる.

反射律 T $\Box\varphi \rightarrow \varphi$

反射律を認めるフレームを仮定すると, 全ての可能世界においてそれぞれの可能世界が自身の可能世界に到達する事が出来る事を示している. このとき, ある可能世界で「到達可能性関係によって繋がっている先全ての可能世界において φ が成り立つ」場合を考えると, それぞれの可能世界は自分自身へその到達可能性関係を持つ為, その可能世界自身で

も $\Box$ 演算子のついた論理式 φ が成り立っている必要がある。また $\Box$ を「知っている」と読んだとき、「 φ を知っている」ならば φ が成り立つ。

推移律 4 $\Box\varphi \rightarrow \Box\Box\varphi$

推移律を認めるフレームにおいては、ある可能世界 s から到達出来る可能世界 t に、そこから到達可能性のある別の可能世界 u が存在する時、可能世界 s から可能世界 u へ到達出来る事を示している。可能世界 t において全ての到達可能性関係のによって繋がった先で φ が成り立つとき、可能世界 t においては $\Box\varphi$ が成り立ち、その可能世界 t に到達可能性関係を持つ可能世界 s においては $\Box\Box\varphi$ が成り立っている必要がある。また $\Box$ を「知っている」と読んだとき、「 φ を知っている」ならば「「 φ を知っている」ということを知っている」が成り立つ。

継続律 D $\Box\varphi \rightarrow \Diamond\varphi$

継続律を認めるフレームにおいては、全ての可能世界において到達可能性関係が存在する事を示している。ある可能世界において $\Box\varphi$ が成り立っているとき、その世界ではある到達可能性で繋がった先の世界で φ が成立する必要がある。

対称律 B $\varphi \rightarrow \Box\Diamond\varphi$

対称律を認めるフレームにおいては、ある可能世界 s から到達出来る可能世界 t に、逆方向にも到達可能性が存在する事を示している。 φ が成り立っている世界から到達可能性関係で繋がったすべての世界から到達可能性関係で繋がったある世界で、 φ が成り立つ必要がある。

ユークリッド律 5 $\Diamond\varphi \rightarrow \Box\Diamond\varphi$

推移律を認めるフレームにおいては、ある可能世界 s から到達出来る可能世界 t と、 s から到達可能性のある別の可能世界 u が存在する時、可能世界 t から可能世界 u へ到達出来る事を示している。ある世界から到達可能性関係で繋がった先の世界で φ が成り立つとき、その世界から到達可能性関係で繋がったすべての世界から到達可能関係で繋がったある世界では φ が成り立つ必要がある。また $\Box$ を「知っている」と読んだとき、「 φ かもしれない」ならば「「 φ かもしれない」ことを知っている」が成り立つ。

到達可能性関係 R とこれらの論理式の間をまとめると以下ようになる。すべてのフレーム $F = (W, R)$ について、

- R が反射的 $\iff$ Tが F で妥当
- R が対称的 $\iff$ Bが F で妥当
- R が推移的 $\iff$ 4が F で妥当

- R がユークリッド的 $\iff 5$ が F で妥当
- R が継続的 $\iff D$ が F で妥当

の同値関係が成立する.

3.2.4 様相論理のヒルベルト流公理系

様相論理のヒルベルト流公理系を定義する. 様相論理のヒルベルト流公理系 HK は以下の公理と推論規則により成り立つ.

- (A1) $\varphi \rightarrow (\psi \rightarrow \varphi)$
- (A2) $(\varphi \rightarrow (\psi \rightarrow \theta)) \rightarrow ((\varphi \rightarrow \psi) \rightarrow (\varphi \rightarrow \theta))$
- (A3) $(\neg\varphi \rightarrow \neg\psi) \rightarrow (\varphi \rightarrow \psi)$
- (MP) φ と $\varphi \rightarrow \psi$ から, ψ を導いてもよい
- (K) $\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$
- (Nec) φ から $\Box\varphi$ を導いてもよい

次に様相論理のヒルベルト式公理系 HK の証明を定義する.

- 定義
以下の条件を満たす連続する論理式 $B_1, \dots, B_n$ を HK の証明とする.
 - B_1 は (A1), (A2), (A3), (K) のいずれかであるか,
 - $B_i (1 < i \leq n)$ は (A1), (A2), (A3), (K) のいずれかであるか, $B_j, B_k (j, k < i)$ から (MP) によって導き出された式であるか, $B_j (j < i)$ から (Nec) によって導き出された式である.

次に先に挙げた表 3.2 の代表的な論理式とヒルベルト流公理系を用いて新しい公理系を以下のように定義する.

- HKD (A1),(A2),(A3),(K),(D),(MP),(Nec)
- HKT (A1),(A2),(A3),(K),(T),(MP),(Nec)
- HKB (A1),(A2),(A3),(K),(B),(MP),(Nec)
- HS4 (A1),(A2),(A3),(K),(T),(4),(MP),(Nec)
- HS5 (A1),(A2),(A3),(K),(T),(5),(MP),(Nec)

これらの公理系はHKと同じように証明と定理の概念を定義することができる。 φ がクリプキモデル M で妥当であるとは φ がすべての $w \in W$ について $M, w \models \varphi$ となることであり、このとき以下の健全性、完全性が成立することが知られている。

- $\vdash_{\text{HK}} \varphi \iff$ すべてのクリプキモデル M において φ は妥当である。
- $\vdash_{\text{HKD}} \varphi \iff$ すべての継続的なクリプキモデルにおいて φ は妥当である。
- $\vdash_{\text{HKT}} \varphi \iff$ すべての反射的なクリプキモデル M において φ は妥当である。
- $\vdash_{\text{HKB}} \varphi \iff$ すべての対称的なクリプキモデル M において φ は妥当である。
- $\vdash_{\text{HS4}} \varphi \iff$ すべての反射的でかつ、推移的なクリプキモデル M において φ は妥当である。
- $\vdash_{\text{HS5}} \varphi \iff$ すべての反射的でかつ、対称的でかつ、ユークリッド的なクリプキモデル M において φ は妥当である。

3.2.5 自然演繹

推論規則

様相論理の自然演繹の推論規則は以下の通りである。また、 $i : A$ と irj における i, j は自然数である。

$$\begin{array}{c}
\frac{i : \varphi_1 \wedge \varphi_2}{i : \varphi_i} (\wedge E) \quad \frac{i : \varphi_1 \quad i : \varphi_2}{i : \varphi_1 \wedge \varphi_2} (\wedge I) \quad \frac{i : \varphi \vee \psi \quad \begin{array}{c} [i : \varphi]_a \\ \vdots \\ \mathbf{k} : \chi \end{array} \quad \begin{array}{c} [i : \psi]_b \\ \vdots \\ \mathbf{k} : \chi \end{array}}{\mathbf{k} : \chi} (\vee E)_{a,b} \quad \frac{i : \varphi_i}{i : \varphi_1 \vee \varphi_2} (\vee I) \\
\\
\frac{i : \varphi \rightarrow \psi \quad i : \varphi}{i : \psi} (\rightarrow E) \quad \frac{\begin{array}{c} [i : \varphi]_a \\ \vdots \\ i : \psi \end{array}}{i : \varphi \rightarrow \psi} (\rightarrow I)_a \quad \frac{i : \neg \varphi \quad i : \varphi}{i : \perp} (\neg E) \quad \frac{\begin{array}{c} [i : \varphi]_a \\ \vdots \\ i : \perp \end{array}}{i : \neg \varphi} (\neg I)_a \\
\\
\frac{i : \perp}{\mathbf{k} : \psi} (\perp) \quad \frac{\begin{array}{c} [i : \neg \varphi]_a \\ \vdots \\ i : \perp \end{array}}{i : \varphi} (RAA)_a
\end{array}$$

様相演算子についての推論規則を加える。

$$\frac{\begin{array}{c} [irj]_a \\ \vdots \\ j : \perp \end{array}}{i : \Box \varphi} (\Box I)_{a\dagger} \quad \frac{i : \Box \varphi \quad irj}{j : \varphi} (\Box E) \quad \frac{j : \Diamond \varphi \quad irj}{i : \Diamond \varphi} (\Diamond I) \quad \frac{i : \Diamond \varphi \quad \begin{array}{c} [irj]_a [j : \varphi]_b \\ \vdots \\ \mathbf{k} : \psi \end{array}}{\mathbf{k} : \psi} (\Diamond E)_{a,b\dagger}$$

$\dagger : j$ は irj を除く $j : \varphi$ を導くために使ったどの仮定にも出現しない.

$\ddagger : j$ は以下の性質を満たす: 1) $i \neq j$, 2) $j \neq k$, 3) j は irj と $j : \varphi$ を除く, $k : \psi$ を導くために使ったどの仮定にも出現しない.

表 3.2 の五つの公理, **T**, **4**, **D**, **B**, **5** をそれぞれ規則 ρ , σ , τ , ϕ , η とすると次のように表すことができる.

$$\frac{}{iri} (\rho) \quad \frac{irj}{jri} (\sigma) \quad \frac{irj \quad irk}{irk} (\tau) \quad \frac{irj \quad irk}{jrk} (\phi) \quad \frac{\begin{array}{c} [irj]_a \\ \vdots \\ k : \psi \end{array}}{k : \psi} (\eta)_a^*$$

$*$: j は以下の性質を満たす: 1) $i \neq j$, 2) $j \neq k$, 3) j は irj を除く, $k : B$ を導くために使ったどの仮定にも出現しない.

ここで $[i : \varphi]$ または $[irj]$ のように各括弧がついている論理式は仮定であり, 推論規則の適用後に仮定のリストがら消される. また様相論理の自然演繹で式 φ が定理になるとは, 仮定のリストが空になり, φ に至る証明図が存在することである.

例 3.3.2

公理系 S4 において $\Box p \rightarrow \Box \Box p$ に至る証明図は次のように与えられる.

$$\frac{\frac{\frac{[0 : \Box p] \quad \frac{[0r1]_b \quad [1r2]_c}{0r2} (\tau)}{2 : p} (\Box E)}{\frac{1 : \Box p}{1 : \Box p} (\Box I)_{c,2:\text{fresh}}}}{\frac{0 : \Box \Box p}{0 : \Box \Box p} (\Box I)_{b,1:\text{fresh}}}}{\frac{0 : \Box p \rightarrow \Box \Box p}{0 : \Box p \rightarrow \Box \Box p} (\rightarrow I)_a}$$

様相論理の意味論と自然演繹の間には完全性・健全性が成り立っていることが知られており, すべての式 φ に対して, 以下の同値関係が成り立つ.

- $\vdash_{\text{HK}} \varphi \iff 0 : \varphi$ が追加の規則を用いなくて様相論理の自然演繹で証明できる
- $\vdash_{\text{HKD}} \varphi \iff 0 : \varphi$ が規則 η を用いて様相論理の自然演繹で証明できる
- $\vdash_{\text{HKT}} \varphi \iff 0 : \varphi$ が規則 ρ を用いて様相論理の自然演繹で証明できる
- $\vdash_{\text{HKB}} \varphi \iff 0 : \varphi$ が規則 σ を用いて様相論理の自然演繹で証明できる
- $\vdash_{\text{HS4}} \varphi \iff 0 : \varphi$ が規則 ρ, τ を用いて様相論理の自然演繹で証明できる
- $\vdash_{\text{HS5}} \varphi \iff 0 : \varphi$ が規則 ρ, σ, τ を用いて様相論理の自然演繹で証明できる

3.3 公開告知論理

公開告知論理とは、様相論理における必然性演算子を「 $\dots$ を知っている」という意味で捉え、信念について記述できる認識論理を拡張させた動的認識論理 (Dynamic epistemic logic) の一種である。認識論理では、認識演算子 K (Knowledge operator) を定義し、様相論理で用いられる必然性演算子 $\Box\varphi$ を $K_a\varphi$ として書き換える。 $K_a\varphi$ と記述することにより、「エージェント a が φ と知っている」と解釈される。公開告知論理では、その認識論理の認識演算子に加え、告知演算子 $[]$ により、モデルの更新を可能にし、エージェントの信念変化について記述することができる。

3.3.1 構文論

命題とエージェント

公開告知論理において命題集合と、エージェント集合の二つの集合が定義されている必要がある。

- 命題集合 $p, q, r, \dots \in P$
- エージェント集合 $a, b, c, \dots \in A$

認識演算子

- 認識演算子 $K : K_a\varphi$ 「エージェント a は φ であることを知っている。」
- 共通知識 (common knowledge) : $C_B\varphi$ 「グループ B のすべてのエージェントは φ が共通知識である。」

共通知識とは、互いがその対象について知っていて、お互いが「相手が知っている」ということも知っている。さらに「相手が自分が知っている」ことも知っているというように、際限なく続くときの知識のことをいう。

告知演算子

- 告知演算子 $[] : []\varphi\psi$ 「 φ という告知の後で ψ である」

告知演算子により、告知により変化したモデルでの推論を可能にする。

言語

命題変数の集合を P , エージェントの集合を A とすると, 言語は $\mathcal{L}_{K[]}(A, P)$ または $\mathcal{L}_{KC[]}(A, P)$ と表される. $\mathcal{L}_{KC[]}(A, P)$ は共通知識も記述できる言語である. ここでの言語とは, 公開告知論理で記述するために必要となる記号の集合を意味している.

$\mathcal{L}_{K[]}(A, P)$ を BNF 記法で書くと,

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi \mid [\varphi]\psi$$

また $\mathcal{L}_{KC[]}(A, P)$ を BNF 記法で書くと, $B \subseteq A$ とし,

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi \mid C_B\varphi \mid [\varphi]\psi$$

$\varphi \vee \varphi$ と $\varphi \rightarrow \varphi$ は $\varphi \wedge \varphi$ と $\neg\varphi$ によって定義できるので省略できる.

3.3.2 意味論

公開告知論理のクリプキモデル $M = (W, (R_a)_{a \in A}, V)$ は以下のように定義される.

- W は可能世界の集合である.
- $R_a \subseteq W \times W$ はエージェント a の到達可能性関係で反射的, 対称的, 推移的, すなわち同値関係である.
- $V : P \rightarrow \wp(W)$ は付値関数である. ($\wp$ とはべき集合である.)

また, クリプキモデル $M = (W, (R_a)_{a \in A}, V)$ に対しての真理条件を次のように帰納的に定義できる. また今回は共通知識について扱わないので言語 $\mathcal{L}_{K[]}(A, P)$ についてのみ記述する.

1. $M, w \models p \iff w \in V(p)$
2. $M, w \models \varphi \wedge \psi \iff M, w \models \varphi$ かつ $M, w \models \psi$
3. $M, w \models \neg\varphi \iff M, w \not\models \varphi$
4. $M, w \models K_a\varphi \iff$ すべての $t \in W : wR_at$ に対し $M, t \models \varphi$
5. $M, w \models [\varphi]\psi \iff M, w \models \varphi$ ならば $M|_{\varphi}, w \models \psi$

ここで、変更後のクリプキモデルである $M|\varphi = (W', (R'_a)_{a \in A}, V')$ は以下のように定められる。

$$\begin{aligned} W' &= |\varphi|_M \\ R'_a &= R_a \cap (|\varphi|_M \times |\varphi|_M) \\ V'(p) &= V(p) \cap |\varphi|_M \end{aligned}$$

また $|\varphi|$ とは φ が真になる可能世界を集めた集合であり、 φ の真理集合 (truth set) と呼ばれ、以下のように表される。

$$|\varphi|_M = \{w \in W \mid M, w \models \varphi\}$$

$M|\varphi$ は、 φ が真である可能世界やその可能世界同士の到達可能性関係だけをクリプキモデルの対象としている。言い換えると φ が偽になる可能世界とそれに繋がる到達可能性関係を削除している。このように定義することにより、告知により変化するクリプキモデルの中で推論が可能となる。

公開告知論理における妥当な論理式とは任意のモデル (W, R, V) で真になる論理式である。

公開告知論理の公理系

以下に公開告知論理のヒルベルト流公理系 PA を示す。命題論理の公理に加え以下の公理と推論規則がある。

- $K_a(\varphi \rightarrow \psi) \rightarrow (K_a\varphi \rightarrow K_a\psi)$
- $K_a\varphi \rightarrow \varphi$
- $K_a\varphi \rightarrow K_aK_a\varphi$
- $\neg K_a\varphi \rightarrow K_a\neg K_a\varphi$
- $[\varphi]p \leftrightarrow (\varphi \rightarrow p)$
- $[\varphi]\neg\psi \leftrightarrow ([\varphi]\psi \wedge [\varphi]\chi)$
- $[\varphi]K_a\psi \leftrightarrow (\varphi \rightarrow K_a[\varphi]\psi)$
- $[\varphi][\psi]\chi \leftrightarrow [\varphi \wedge [\varphi]\psi]\chi$
- φ と $\varphi \rightarrow \psi$ から ψ が導かれる
- φ より $K_a\varphi$ が導かれる

このとき以下の健全性・完全性が成り立つことが知られてる。

- $\vdash_{PA} \varphi \iff \varphi$ がすべての公開告知論理のクリプキモデルで妥当である。

3.3.3 クリプキモデルの例

公開告知論理によってクリプキモデルがどのように変化していくのかを示す．いまクリプキモデルを以下のように設定する．

- 可能世界の集合 $W = \{w_0, w_1\}$
- 到達可能性関係 $R_a = \{(w_0, w_0), (w_1, w_1)\}$ $R_b = \{(w_0, w_0), (w_0, w_1), (w_1, w_0), (w_1, w_1)\}$
- 付値 $V(p) = \{w_1\}$

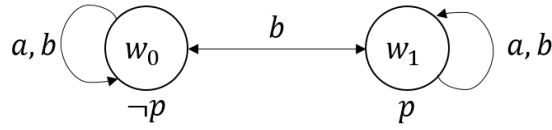


図 3.2 設定したクリプキモデル

上のように設定したときのクリプキモデルが図 3.2 である．設定した到達可能性関係は同値関係を満たしている．このときエージェント b にとって可能世界 w_0 と w_1 にはとつ経つ可能性関係によって繋がれていて，二つは同値関係であり区別がつかない状態にある．エージェント a にとっては w_0 と w_1 には到達可能性関係はないので二つの可能世界は区別がついている．ここで $p \wedge \neg K_b p$ という告知があったとすると，図 3.3 のようにクリプキモデルが変更される．

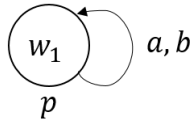


図 3.3 変更後のクリプキモデル

$p \wedge \neg K_b p$ が偽になる可能世界 w_0 とそれに繋がる到達可能性関係は削除され w_1 のみが残る．

3.3.4 誕生日の謎

本研究の解析対象として挙げた sum and product や“誕生日の謎”を公開告知論理を用いて解くことができる．以下では“誕生日の謎”について解法を示していく．2 章にて挙げた引用をもう一度載せる．

A さんと B さんは C さんの誕生日にプレゼントをあげたいのですが，二人とも C さんの誕生日の日付がわかりません．ただ C さんの誕生日は下の 10 組の中のどれかだけわかります．

- 3 月 4 日
- 3 月 5 日
- 3 月 8 日
- 6 月 4 日
- 6 月 7 日
- 9 月 1 日
- 9 月 5 日
- 12 月 1 日
- 12 月 2 日
- 12 月 8 日

C さんは A さんに「何月」誕生日か教えました，そして B さんに「何日」誕生日か教えました．

- C さん「さあ，私の誕生日はいつですか？」
- 1. A さん「俺は答えがわからないから，B さんもわからないはずです」
- 2. B さん「俺もわからなかったけど，今の A さんの言葉でわかりました」
- 3. A さん「あっ，じゃ俺もわかりました」

では C さんの誕生日は 10 組の中のどの日でしょうか？

クリプキモデルの設定

まず可能世界 w にあたるのは誕生日の候補日である．そして誕生日の候補日の集合を $W = [\text{誕生日の候補日 (cob)}]$ とする．

エージェントの集合を A とすると，いま A と B がエージェントであるのでそれぞれ小文字で表現して，

$$A = \{a, b\}$$

次に A と B の到達可能性関係について記述する．A にとって教えられた月が 6 月であれば 6 月の 4 日なのか 7 日なのかが分からない．よって区別できないのは候補日の中で月が同じである日付であるから，同じ月の候補日同士が到達可能性関係で結ばれている．また B にとっては何日かしか分かっていないので，同じ日同士，例えば 3 月 4 日と 6 月 4 日に到達可能性関係が結ばれている．図で表すと図 3.4 のようになる．

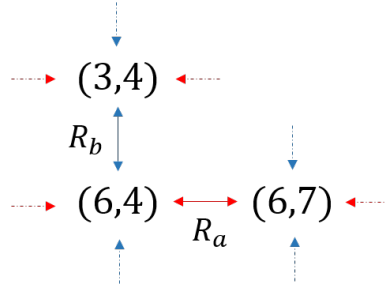


図 3.4 到達可能性関係

赤い矢印はエージェント a の到達可能性関係，青い矢印がエージェント b の到達可能性関係である．またこの図では自分と自分を繋いだ到達可能性関係については省略しているが，本来はそこにも到達可能性関係があり，反射律を満たしている．

先に定義した可能世界とエージェントを用いて両エージェントの到達可能性関係を表すと

$$R_a = \{((x, y), (x', y')) \in W \times W \mid x = x'\}$$

$$R_b = \{((x, y), (x', y')) \in W \times W \mid y = y'\}$$

と書くことができる．

次に命題と付値を与える．可能世界 $(6,7)$ であれば「6月7日はCの誕生日である」という文が真になり，そしてこの命題が成り立つのも可能世界 $(6,7)$ のみとなる．このときの「6月7日はCの誕生日である」という一文が一つの命題になる．そして付値は $V(\text{「6月7日はCの誕生日である」}) = \{(6,7)\}$ というふうに記述することができる．これを形式化すると，

$$V(x/y \text{ is the } C\text{'s birthday}) = \{(x, y)\}$$

となり，一意に示せる．論理式中では命題を“ x/y is the C 's birthday”と英語で記述する．

発言の論理式化

命題を元に「AはCの誕生日を知っている」を論理式で表す．Aは誕生日を知っているのだが，それが何月何日かという情報はなく，「AはCの誕生日が3月4日か3月5日か3月8日か…を知っている」というふうに意識することが可能である．論理式で表すと，「エージェント $z \in \mathcal{A}$ はCの誕生日を知っている」は，

$$\bigvee_{(x,y) \in [Cob]} K_z(x/y \text{ is the } C\text{'s birthday})$$

また「エージェント $z \in \mathcal{A}$ はCの誕生日を知らない」は，

$$\bigwedge_{(x,y) \in [Cob]} \neg K_z(x/y \text{ is the } C\text{'s birthday})$$

問題文に戻り，二人の会話を形式化していく．一つ目の A の発言である「俺は答えがわからないから，B くんもわからないはずです」は，「B が誕生日を知らないことを知っていた」というふうに捉えることができる．二つ目の発言は，一つ目の発言を受けて「誕生日をしています」，三つ目の発言は一つ目，二つ目の発言を受けて，「誕生日が知っています」と捉えられる．よって 3 つの発言を形式化すると，

1. A 「B は C の誕生日がわからないはずです」

$$\text{ann1} = K_a(\bigwedge_{(x,y) \in [Cob]} \neg K_b(x/y \text{ is the } C\text{'s birthday}))$$
2. B 「今の A の言葉で C の誕生日がわかりました」

$$\text{ann2} = [\text{ann1}](\bigvee_{(x,y) \in [Cob]} K_b(x/y \text{ is the } C\text{'s birthday}))$$
3. A 「私も C の誕生日がわかりました」

$$\text{ann3} = [\text{ann1}][\text{ann2}](\bigvee_{(x,y) \in [Cob]} K_a(x/y \text{ is the } C\text{'s birthday}))$$

となる．二つ目の告知，三つ目の告知は前の告知により変化したクリプキモデル上での告知になるので，告知演算子 $[]$ を用いて表される．

告知によるモデルの変化

形式化したクリプキモデルが告知により変化していく様子について考察していく．

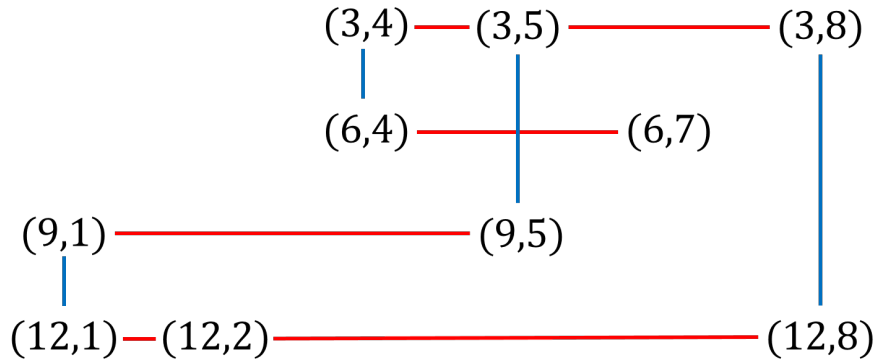


図 3.5 初期状態

図 3.5 が C の誕生日の候補日を世界の集合とし，その間における到達性関係を記した初期状態の図になる．ただし自分自身への到達可能性関係はカットしている．

一つ目の告知によりクリプキモデルは変化し，図 3.6 になる．「B にとって区別がつく世界，つまりある可能世界において他の可能世界と B の到達可能性関係によって繋がれた世界がないような可能世界」と A の到達可能性関係によって繋がれている世界，つまり 6 月の候補日と 12 月の候補日は全て告知が成り立たないので削除される．また到達可能性

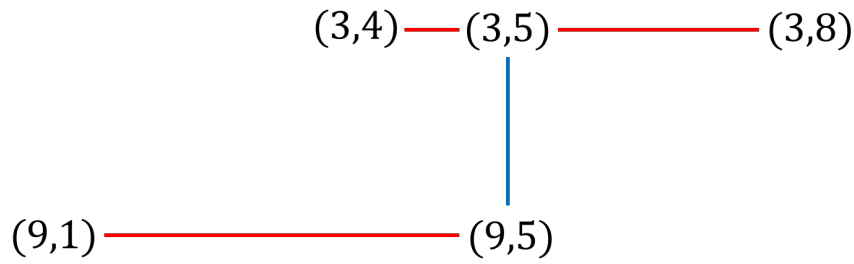


図 3.6 一つ目の告知により変化した状態

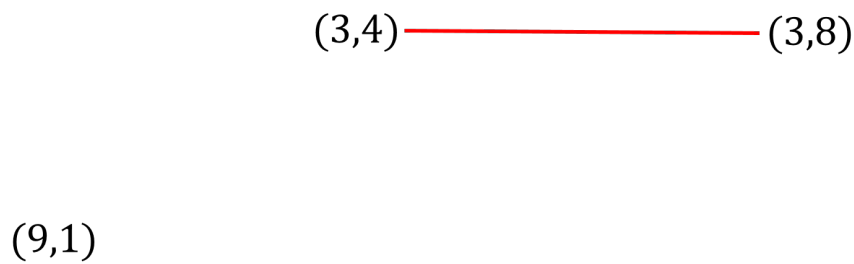


図 3.7 二つ目の告知により変化した状態

関係は元のモデルの到達可能性関係を維持しつつ、消えた可能世界への到達可能性関係は削除される。

二つ目の告知により変化したクリプキモデルが図 3.7 になる。これは一つ目の告知により変化したモデル上で、B にとっていまだ区別がついていない可能世界である 3 月 5 日と 9 月 5 日が削除される。

(9,1)

図 3.8 三つ目の告知により変化した状態

そして最後の告知を受けて生成されるクリプキモデルが図 3.8 になる。最後に可能世界 (9,1) だけが残る、この可能世界だけがすべての告知を満たしていることが分かり、このパズルの解である C の誕生日は 9 月 1 日であったことが分かる。

このようにして論理パズルをエージェントの信念変化と捉えることで、第二章の人間の推論と同じような思考の推移を辿り、動的にシミュレートすることができる。

第4章 Inquisitive semantics

Ivano, Jeroen, Floris らによる Inquisitive semantics [3] とは、平叙文と疑問文の両方に統一的意味を与える形式意味論であり、本章では Inquisitive semantics の概念を動的認識論理に応用した Ivano, Floris の Inquisitive dynamic epistemic logic (IDEL) [8] について説明していく。IDEL も用いることで到達可能性関係に違った尺度で推論することを可能にし、エージェントが「疑っている状態」を表現することができる。まず Inquisitive semantics における重要な概念である issue について記述し、IDEL の構文論、意味論を与え、IDEL を用いて行うことができる推論について記述する。

4.1 Information states と issue

公開告知論理においては告知により可能世界を削られていくことに情報量があった。 $p \vee \neg p$ や $q \vee \neg q$ のような論理式を告知したとしても公開告知論理において論理式はトートロジーであり、すべての可能世界で成り立つので可能世界の数は変わることはなく、この論理式には情報量はない。しかし $p \vee \neg p$ と $q \vee \neg q$ は全く別の告知であり、そこに何らかの差異が記述できるように思われる。少なくとも $p \vee \neg p$ という告知があったなら p に関心を抱いているのだということが分かる。その関心を抱く部分について形式化することができるのが IDEL の特徴である。

4.1.1 Information states

この informativeness と inquisitiveness を表現することに必要な概念が information states と Issue である。Information state とは可能世界の集合を W としたときの部分集合であり、information state $s \subseteq W$ で表される。可能世界の集合 W において何か情報が与えられたときに絞り込まれてできる可能世界の集合である。

極端な例としては information state が空集合であるとき、これを矛盾状態 (inconsistent state) と呼び、information state が可能世界の集合自身であるとき、何も情報をもっていないので、無知状態 (ignorant state) と呼ぶ。

図 4.1 は可能世界の集合 $W = \{w_1, w_2, w_3, w_4\}$ としたときの information state の例である。右から左に向かうにつれて情報が多くなっていき、information state に含まれる可能世界の数はだんだんと少なくなっている。

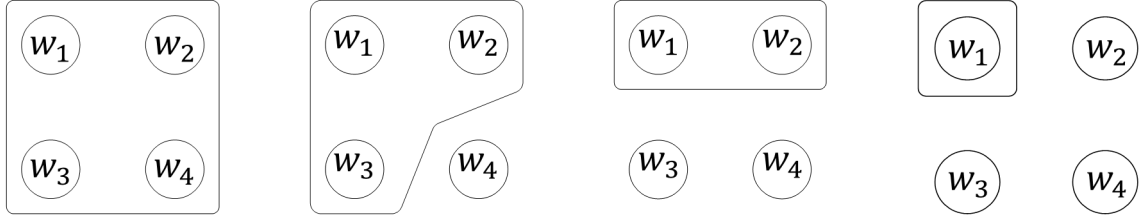


図 4.1 Information states

4.1.2 Issues

Issue とは知りたい情報に対して，その情報の意味を区別する尺度である．ある文脈がなんらかの情報を持っている場合，それは information state s として表現され，更に詳しい内容を判別するためにその s の中でグループ分けを行う． W を可能世界の集合，information state s を W の部分集合とすると，issue I は空でない集合であり，以下のような性質を持っている．

1. I は downward closed である： $t \in I$ かつ $t' \subseteq t$ ならばまた $t' \in I$
2. I は s の issue である： $\bigcup I = s$

downward closure という性質は issue がその内部に密に詰まっていることを表現している．

また information state s に対するすべての issue を Π_s と表し，可能世界の集合 W に対する，すべての information state に対する issue の集合を $\Pi = \bigcup_{s \subseteq W} \Pi_s$ として表す．

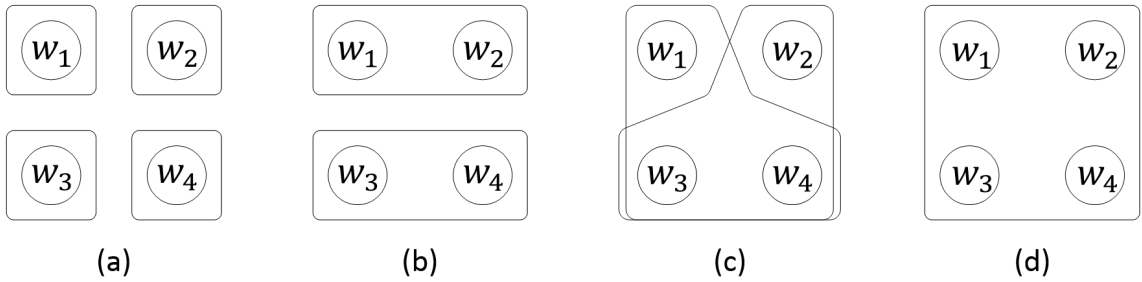


図 4.2 Issues

図 4.2 は information state $s = \{w_1, w_2, w_3, w_4\}$ に対する 4 つの issue である．この図は issue の中で最大の要素のみを描画している．(a) ではこれ以上 s を分解をわけることができない issue であり， $I = \{\{w_1\}, \{w_2\}, \{w_3\}, \{w_4\}\}$ となる．(b) は s に対して s を $\{w_1, w_2\}$ と $\{w_3, w_4\}$ とにグループが分かれている．このときの issue は $I = \{\{w_1, w_2\}, \{w_1\}, \{w_2\}, \{w_3, w_4\}, \{w_3\}, \{w_4\}\}$ になる．(c) では $\{w_1, w_3, w_4\}$ と $\{w_2, w_3, w_4\}$ のグループに分けら

れる．これは可能世界が同時に二つのグループに入っているという issue の興味深い性質を示している．最後に (d) は s に対して自明となる issue である．

この information states と issue の関係を到達可能性関係に用いる．Information state を DEL におけるエージェント a の可能世界 w における到達可能性関係とし，これを $\sigma_a(w)$ とする．また inquisitive state として $\Sigma_a(w)$ を導入する．これは $\sigma_a(w)$ に対する issue である．これにより各エージェントは到達可能性関係によってつながれている可能世界に対してより正確に表現できるようになる．このとき二つの関係は $\sigma_a(w) = \bigcup \Sigma_a(w)$ である．

4.2 Inquisitive Dynamic Epistemic Logic

構文論を定義し，その後 $\Sigma_a(w)$ を用いて IDEL における意味論を定義していく．

4.2.1 構文論

公開告知論理と同様に命題集合を $p, q, r, \dots \in P$ ，エージェント集合を $a, b, c, \dots \in A$ と前提として与えておく．

クリプキモデル

$\Sigma_a(w)$ を用いてクリプキモデルがどのように表されるかを記述する．クリプキモデルを $M = (W, (\Sigma_a)_{a \in A}, V)$ とすると

- W は可能世界の集合である．
- $(\Sigma_a)_{a \in A} : W \rightarrow \Pi$ となる写像であり，ここで可能世界 w からの写像である $\Sigma_a(w)$ には以下の性質がある．
 - Factivity : すべての $w \in W$ に対し， $w \in \sigma_a(w)$
 - Introspection : すべての $w, v \in W$ に対し， $v \in \sigma_a(w)$ ならば $\Sigma_a(v) = \Sigma_a(w)$

$\sigma_a(w) := \bigcup \Sigma_a(w)$ はエージェント a の w における information state を表している

- $V : P \rightarrow \wp(W)$ は付値関数である．

言語

IDEL では平叙文と疑問文を同時に扱える構文が必要である．また疑問文にはいくつか制約がある．例えば平叙文と平叙文の連言をとったときにできる文は平叙文であるし，疑問文と疑問文の連言をとったときできた文は疑問文でなくてはならない．そして平叙文と疑問文との連言をとることはできない．これは直感的に考えてみても「彼は犯人である」と「凶器は包丁ですか?」という2文の和をとってできるような「彼は犯人であり，かつ凶器は包丁ですか?」という文章はおかしいことが分かる．また「ならば」で2文を繋いだとき，前件である文は平叙文でなくてはならず，合成してできた文は後件に依存する．「証拠がありますか?」と「彼は犯人である」という2文を「ならば」で繋ぐと，「証拠がありますか?ならば彼は犯人である」というふうに全く意味不明な文章になってしまう．以上を踏まえて構文を構築する．

平叙文全体の集合を $\mathcal{L}_!$ ，疑問文全体の集合を $\mathcal{L}_?$ とし，疑問文 $\alpha, \beta, \gamma \in \mathcal{L}_!$ と平叙文 $\mu, \nu, \lambda \in \mathcal{L}_?$ と限定しない文 $\varphi, \psi, \chi \in \mathcal{L}_! \cup \mathcal{L}_?$ を定義すると，以下に示す構文となる．

1. すべての $p \in P$ に対し $p \in \mathcal{L}_!$
2. $\perp \in \mathcal{L}_!$
3. $\alpha_1, \dots, \alpha_n \in \mathcal{L}_!$ ならば $?\{\alpha_1, \dots, \alpha_n\} \in \mathcal{L}_?$
4. $\varphi \in \mathcal{L}_\circ$ かつ $\psi \in \mathcal{L}_\circ$ ならば $\varphi \wedge \psi \in \mathcal{L}_\circ$ ，ここで $\circ \in \{!, ?\}$
5. $\alpha \in \mathcal{L}_!$ かつ $\varphi \in \mathcal{L}_\circ$ ならば $\alpha \rightarrow \varphi \in \mathcal{L}_\circ$ ，ここで $\circ \in \{!, ?\}$
6. $\circ \in \{!, ?\}$ であり， $\varphi \in \mathcal{L}_\circ$ かつ $a \in A$ ならば $K_a \varphi \in \mathcal{L}_!$
7. $\circ \in \{!, ?\}$ であり， $\varphi \in \mathcal{L}_\circ$ かつ $a \in A$ ならば $E_a \varphi \in \mathcal{L}_!$
8. $\mathcal{L}_!$ にも $\mathcal{L}_?$ にも属していない文はない

3つ目の項目について説明する．論理式 $?\{\alpha_1, \dots, \alpha_n\}$ は平叙文 $\alpha_1, \dots, \alpha_n$ から成り立っている．平叙文により選択肢を与え，その選択肢のうちに当てはまるものがありますか? というような疑問文になる．例えば「犯人はAですか?Bですか?それともCですか?」という文はこの論理式を用いて， $?\{A \text{ が犯人}, B \text{ が犯人}, C \text{ が犯人}\}$ というふうに表される．さらに， $\alpha_1, \dots, \alpha_n$ は原始命題ではなくてよく，つまり論理式には疑問文の性質を補成分として持っていることになる．

4, 5項目は疑問文の制約である．そして論理演算子の「否定」と「または」は「かつ」と「ならば」を用いて定義される．平叙文 α と β を用いて，「否定」は $\neg\alpha = \alpha \rightarrow \perp$ ，「または」は $\alpha \vee \beta = \neg(\neg\alpha \wedge \neg\beta)$ として意義付けされる．また「否定」を導く際，「ならば」を使用しているので「否定」の演算子を用いることができるのは平叙文の場合のみとなる．「または」も同様に「否定」を用いて定義されているので，平叙文にしか使うことが

できない．さらに，疑問文の基本形である $\{\alpha, \neg\alpha\}$ は Yes/No 疑問文と呼ばれ， $?\alpha$ として表す．

6 項目目は認識演算子 K_a であり，エージェントの情報を記述することができる．7 項目目の E_a はエージェントが抱えている (Entertain する) issue について記述することができる考慮演算子である．この二つは平叙文，疑問文の両方に対して用いることが可能であり，用いてできた論理式はすべて平叙文となる．例えば $K_a?p$ という論理式は「 p かどうか知っている」という意味の平叙文となる．これは「 p を知っている，または $\neg p$ を知っている」と同じ意味になり， $K_ap \vee K_a\neg p$ として書くこともできる．この二つの演算子を用いることで，「エージェント a は p かどうか疑っている」状態というのを記述することができる．

4.2.2 意味論

様相論理や公開告知論理の意味論というのはクリプキモデルに対して，可能世界と論理式との間の二項関係として再帰的に定義されてきた．しかし疑問文を扱うにあたり，可能世界と論理式の関係では不都合なことが起こる．疑問文は与えられた可能世界で真かどうかというこは分からないのだが，何の情報がその疑問文に対して答えるのに必要かということはある．そのため，疑問文を評価するのは可能世界に対してではなく，可能世界の集合である information state に対してであり，そのときの疑問文と可能世界の集合の間関係は，真理条件というよりは，解明条件 (resolution condition) であるといえる．

この分析状態を用いて疑問文を表し，平叙文を真なる状態として意味論を定義できるのだが，平叙文に対しても information state を用いて表現することができ，統一的な意味論を作ることができる．

解明条件

M を IDEL のクリプキモデル， s を M の information state とすると，解明条件は以下のように与えられる．

1. $M, s \models p \iff$ すべての可能世界 $w \in s$ に対して $w \in V(p)$
2. $M, s \models \perp \iff s = \emptyset$
3. $M, s \models ?\{\alpha_1, \dots, \alpha_n\} \iff$ ある $1 \leq i \leq n$ に対して $M, s \models \alpha_i$
4. $M, s \models \varphi \wedge \psi \iff M, s \models \varphi$ かつ $M, s \models \psi$
5. $M, s \models \alpha \rightarrow \varphi \iff$ すべての $t \subseteq s$ に対し， $M, t \models \alpha$ ならば $M, t \models \varphi$
6. $M, s \models K_a\varphi \iff$ すべての $w \in s$ に対し $M, \sigma_a(w) \models \varphi$

7. $M, s \models E_a \varphi \iff$ すべての $w \in s$ かつ, すべての $t \in \Sigma_a(w)$ に対して $M, t \models \varphi$

これに二つの事実が挙げられる.

- Fact1 : $M, s \models \varphi$ かつ $t \subseteq s$ ならば $M, t \models \varphi$ (downward closure)
- Fact2 : すべてのモデル M , すべての論理式 φ に対し $M, \emptyset \models \varphi$

一つ目の事実として, ある information state においてある論理式が真であるとき, その部分集合においてもある論理式が真であるという持続性の性質を持っているものである. 二つ目の事実, information state が空集合である場合, 全ての論理式が成り立つというものである.

また平叙文に関しては, 以下のような書き換えが成り立つ.

- Fact3 : $M, s \models \alpha \iff$ すべての $w \in s$ に対して $M, \{w\} \models \alpha$

これは $(\Rightarrow)$ は Fact1 により導かれ, $(\Leftarrow)$ は α の構成に関する帰納法によって導かれる.

「否定」と「または」の演算子を用いた論理式と Yes/No 疑問文は次のように定義される. またこれは Ivano, Floris[8] の論文中には明記されていなかったので証明を加えた.

否定

$$\begin{aligned}
 M, s \models \neg \alpha &\iff M, s \models \alpha \rightarrow \perp \\
 &\iff \text{すべての集合 } t \subseteq s \text{ に対し, } M, t \models \alpha \text{ ならば } M, t \models \perp \\
 &\iff \text{すべての集合 } t \subseteq s \text{ に対し, } M, t \models \alpha \text{ ならば } t = \emptyset \\
 &\iff \text{すべての集合 } t \subseteq s \text{ に対し, } t \neq \emptyset \text{ ならば } M, t \not\models \alpha \\
 &\iff \text{すべての空でない集合 } t \subseteq s \text{ に対し, } M, t \not\models \alpha
 \end{aligned}$$

または

$$\begin{aligned}
 M, s \models \alpha \vee \beta &\iff M, s \models \neg(\neg \alpha \wedge \neg \beta) \\
 &\iff \text{すべての空でない集合 } t \subseteq s \text{ に対し, } M, t \not\models \neg \alpha \wedge \neg \beta \\
 &\iff \text{すべての空でない集合 } t \subseteq s \text{ に対し, } M, t \not\models \neg \alpha \text{ または } M, t \not\models \neg \beta \\
 &\iff \text{すべての空でない集合 } t \subseteq s \text{ に対し, ある空でない集合 } t' \subseteq t \text{ に対し} \\
 &\quad \underline{M, t' \models \alpha, \text{ かつ, ある空でない集合 } t'' \subseteq t \text{ に対し } M, t'' \models \beta}_{\textcircled{1}} \\
 &\iff \underline{\text{ある } t_1, t_2 \text{ があり, } t_1 \cup t_2 = s \text{ かつ } t_1 \models \alpha \text{ かつ, } t_2 \models \beta}_{\textcircled{2}}
 \end{aligned}$$

① $\iff$ ② を証明する.

② $\Rightarrow$ ① を示す.

1. $s = \emptyset$ の場合, t は空でない集合より明らかである.

2. $s \neq \emptyset$ の場合

② と空でない集合 $t \subseteq s$ を仮定する.

② より, ある $t_1, t_2 \subseteq t$ があり, $M, t_1 \models \alpha$ かつ $M, t_2 \models \beta$

このとき t は以下のように書き換えられる.

$$t = (t_1 \cup t_2) \cap t = (t_1 \cap t) \cup (t_2 \cap t)$$

Fact1($M, s \models \varphi$ かつ $t \subseteq s$ ならば $M, t \models \varphi$) より,

$$M, t \cap t_1 \models \alpha \text{ かつ } M, t \cap t_2 \models \beta$$

が成り立つことが分かる.

(a) $t_1 \cap t \neq \emptyset$ のとき,

空でない集合 $t_1 \cap t \subseteq t$ があり, それに対し $M, t_1 \cap t \models \alpha$ が成り立つので, ① は真である.

(b) $t_1 \cap t = \emptyset$ のとき,

$t = (t_1 \cap t) \cup (t_2 \cap t)$ かつ $t \neq \emptyset$ より, $t_2 \cap t \neq \emptyset$ である.

空でない集合 $t_2 \cap t \subseteq t$ があり, それに対し $M, t_2 \cap t \models \beta$ が成り立ち, ① は真である.

よって, すべての場合で ① は真であり, ② $\Rightarrow$ ① は示された.

① $\Rightarrow$ ② を示す.

1. $s = \emptyset$ の場合

$t_1 = \emptyset, t_2 = \emptyset$ とすると, Fact2(すべてのモデル M , すべての論理式 φ に対し $M, \emptyset \models \varphi$) より, $M, \emptyset \models \alpha$ かつ $M, \emptyset \models \beta$ が成り立ち, 明らかである.

2. $s \neq \emptyset$ の場合

① を仮定すると一つ一つの可能世界の単元集合 $\{w\}$ 全てに対し $M, \{w\} \models \alpha$ または $M, \{w\} \models \beta$ がいえるので, 次のような関数が作られる.

$f(w) : W \rightarrow \mathcal{L}$ は可能世界を与えると, 次のように平叙文を返す関数である.

$$f(w) = \begin{cases} \alpha & M, \{w\} \models \alpha \\ \beta & M, \{w\} \not\models \alpha \text{ かつ } M, \{w\} \models \beta \end{cases}$$

このとき, $\bigcup f^{-1}[\alpha] = t_1, \bigcup f^{-1}[\beta] = t_2$ とすると, Fact3 より $t_1 \cup t_2 = s$ かつ $M, t_1 \models \alpha$ かつ $M, t_2 \models \alpha$ が成り立つので ② は真である.

よって ① $\Rightarrow$ ② は示され, ① $\Longleftrightarrow$ ② は証明された.

Yes/No 疑問文

$$\begin{aligned} M, s \models ?\alpha &\iff M, s \models ?\{\alpha, \neg\alpha\} \\ &\iff M, s \models \alpha \text{ または } M, s \models \neg\alpha \end{aligned}$$

真理条件

真理条件は可能世界に対する論理式との二項関係であり、可能世界 w において φ が成り立っているとはつまり、state $\{w\}$ において φ が成り立っているということとして定める。

$$M, w \models \varphi \stackrel{\text{def}}{\iff} M, \{w\} \models \varphi$$

この単元集合の考えにより、解明条件より以下のような真理条件が導かれる。

1. $M, w \models p \iff w \in V(p)$
2. $M, w \not\models \perp$
3. $M, w \models ?\{\alpha_1, \dots, \alpha_n\} \iff \text{ある } 1 \leq i \leq n \text{ に対して } M, w \models \alpha_i$
4. $M, w \models \varphi \wedge \psi \iff M, w \models \varphi \text{ かつ } M, w \models \psi$
5. $M, w \models \alpha \rightarrow \varphi \iff M, w \not\models \alpha \text{ かつ } M, w \models \varphi$
6. $M, w \models \neg\alpha \iff M, w \not\models \alpha$
7. $M, w \models \alpha \wedge \beta \iff M, w \models \alpha \text{ または } M, w \models \beta$
8. $M, w \models K_a\varphi \iff M, \sigma(w) \models \varphi$
9. $M, w \models E_a\varphi \iff \text{すべての } t \in \Sigma_a(w) \text{ に対して } M, t \models \varphi$

真理条件は解明条件より導かれるものであり、その実は疑問文を含んではいるが、公開告知論理のそれと同じものになることが分かる。そして真理集合 $|\varphi|_M$ もこの真理条件から導かれる。

$$|\varphi|_M := \{w \in W \mid M, w \models \varphi\}$$

今までの様相論理においては可能世界という一つの尺度においてのみ表すことができたが、IDEL においては state というもう一つの尺度がある。そのため、真理集合とは別の、論理式が真となる state の集合を記述する必要がある。それを proposition とする。モデル M における φ の proposition を $\llbracket \varphi \rrbracket_M$ として表すと、

$$\llbracket \varphi \rrbracket_M := \{s \subseteq W \mid M, s \models \varphi\}$$

となる． Fact1,2 より prosotisation は常に空でない集合であり， issue のような downward closed な性質を持っている． また， proposition の和集合をとると， 真理集合になる．

$$|\varphi|_M = \bigcup \llbracket \varphi \rrbracket_M$$

証明は以下の通りである．

$w \in |\varphi|_M$ を仮定する．

$$\begin{aligned} M, w \models \varphi &\iff M, \{w\} \models \varphi \\ &\iff \{w\} \in \llbracket \varphi \rrbracket_M \\ &\implies w \in \bigcup \llbracket \varphi \rrbracket_M \end{aligned}$$

また逆に $w \in \bigcup \llbracket \varphi \rrbracket_M$ と， $M, s \models \varphi$ となる s に対して $w \in s$ となることを仮定すると， $\{w\} \subseteq s$ より，

$$\begin{aligned} M, \{w\} \models \varphi &\iff M, w \models \varphi \\ &\iff w \in |\varphi|_M \end{aligned}$$

様相演算子

認識演算子 $K_a\varphi$ と考慮演算子 $E_a\varphi$ には関係性があり， 全ての文に対して次の式が成り立つ．

$$M, s \models K_a\varphi \text{ ならば } M, s \models E_a\varphi$$

これは， Fact 1 と $\sigma_a(w) = \bigcup \Sigma_a(w)$ の性質により， $M, \sigma_a(w) \models \varphi$ ならば， 全ての $t \in \Sigma_a(w)$ に対して， $M, t \models \varphi$ が成り立つからである．

またすべての平叙文に対しては $K_a\alpha$ と $E_a\alpha$ は同値になる． これは， state s において α が成り立つということは， その s の要素であるすべての可能世界で真でなければならず， α が $t \in \Sigma_a(w)$ で成り立っているとすると， すべての可能世界 $w \in \bigcup \Sigma_a(w) = \sigma_a(w)$ で真であり， 故に $\sigma_a(w)$ においても真となるからである．

このことから E_a は平叙文に対しては K_a と同じ扱いができる． しかし疑問文 μ に対して E_a はその真価を発揮する． $E_a\mu$ が成り立つのは意味論により，「エージェント a のすべての issue において μ が成り立っている」さらにいうと「エージェント a が興味を抱いているすべての state において μ が成り立っている」ととることができる． これは「エージェント a が μ について疑っている (wonder)」というニュアンスに近い． 例えば a はある疑問文 μ を投げかけられたとき， μ について十分な情報をもっているとき， そこでは $K_a\mu$ が成り立ち， $E_a\mu$ も成り立つ． この場合は μ について疑っているとはいえない． 疑っているということがいえるのは， μ が成り立つかどうかという十分な情報をもっていないとき， つまり $\neg K_a\mu$ であり， しかしその情報を知りたいと抱いている状態， つまり $E_a\mu$ が成り立っている場合である． まとめると， a が μ を疑っているとは μ について知らないが， 知りたがっ

ている状態のことであり、疑っていることを表現する懐疑演算子 (wonder modality) W_a は以下のように表すことができる.

$$W_a\varphi := \neg K_a\varphi \wedge E_a\varphi$$

この三つの演算子について例を与えて実際にクリプキモデル上でどのように振舞うのか見ていく. 以下のように設定したクリプキモデルを与える.

- 可能世界 w の集合 $W = \{w_1, w_2, w_3\}$
- issue $\Sigma_a(w_1) = \Sigma_a(w_2) = \Sigma_a(w_3) = \{\{w_1, w_2\}, \{w_1, w_3\}, \{w_1\}, \{w_2\}, \{w_3\}\}$
- 付値 $V(p) = \{w_1, w_2\}$ $V(q) = \{w_1, w_3\}$

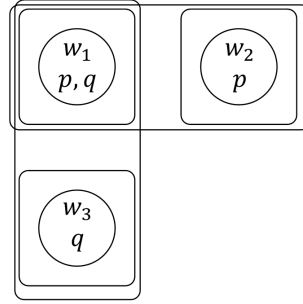


図 4.3 issue の初期状態

図 4.3 は設定したクリプキモデルの各可能世界におけるエージェント a の issue である. このとき, Introspection の性質により, w_1 からみた issue も, w_2 からみた issue も, w_3 からみた issue もすべて同一の issue であり, 和集合をとって一つにしたものも同じ issue になり, 一つの図として表される.

この設定したモデルは大きく分けて二つの性質を持った issue があり, p が成り立っている集合 $\{w_1, w_2\}$ と, q が成り立っている集合 $\{w_1, w_3\}$ とに分けられている. このとき以下のことが可能世界の集合 $s = \{w_1, w_2, w_3\}$ についていえる.

- $M, s \not\models K_a?\{p, q\}$
- $M, s \models E_a?\{p, q\}$

エージェント a にとって w_1 からみたときに, 区別がつかない可能世界は w_1, w_2, w_3 , であり, その全てにおいて, p が成り立つ, もしくは q が成り立つということがいえないので「エージェント a は p か q であることを知っている」は s においてなりたたないの, $M, s \not\models K_a?\{p, q\}$ であることがいえる. また, エージェント a の w_1 における issue についてみたときに, すべての issue は p が成り立つか, または q が成り立つことを満たして

いる．よって「エージェント a は p か q であることを知りたいと抱いている」ということがいえ，同様に w_2, w_3 の可能世界においても成り立つので $M, s \models E_a?\{p, q\}$ は真となる．さらにこのとき $M, s \models \neg K_a?p, q$ も成り立つので， $M, s \models \neg K_a?p, q \wedge E_a?p, q$ が成り立つので，

- $M, s \models W_a?\{p, q\}$

が成り立ち，エージェント a はどの可能世界からみても， p であるか q であるか疑っているといえる．

4.2.3 告知演算子によるモデルの書き換え

モデルの変更のために告知演算子を構文論に加える．

- $\circ \in \{!, ?\}$ であり， $\varphi \in \mathcal{L}$ 。かつ $\bullet \in \{!, ?\}$ であり， $\psi \in \mathcal{L}$ 。ならば $[\varphi]\psi \in \mathcal{L}$ 。

公開告知 $[\varphi]\psi$ の解明条件は，

- $M, s \models [\varphi]\psi \iff M, s \models \varphi$ ならば $M^\varphi, s \models \psi$

また公開告知 $[\varphi]\psi$ の真理条件は

- $M, w \models [\varphi]\psi \iff M, w \models \varphi$ ならば $M^\varphi, w \models \psi$

であり，ここで $M^\varphi = (W^\varphi, (\Sigma_a)_{a \in A}^\varphi, V^\varphi)$ は以下のように定義される．

- $W^\varphi = W \cap |\varphi|_M$
- $(\Sigma_a)_{a \in A}^\varphi = (\Sigma_a)_{a \in A} \cap \llbracket \varphi \rrbracket$
- $V^\varphi(p) = V(p) \cap |\varphi|_M$

φ という告知がなされたとき，可能世界の集合は真理集合 $|\varphi|_M$ に制限され，inquisitive state $\Sigma_a(w)$ は proposition $\llbracket \varphi \rrbracket_M$ によって制限される．

平叙文を告知した場合と，疑問文によって質問された場合では，モデルに対する効果というのも変わってくる．以下に例を示しながら説明する．

図 4.4 は可能世界の集合を $W = \{11, 10, 01, 00\}$ であり，すべての可能世界 w におけるエージェント a の inquisitive state $\Sigma_a(w)$ の最大となる要素を表したものであり，簡単化のため Σ_a と表記する．11 において p も q も真であり，10 では p が真 q が偽，01 では p が偽 q が真，00 では p も q も偽である．図 (a) において， Σ_a は W に対する自明な issue を体現していて， p, q に対して何も情報を持っていないし，何の issue も持っていないことを意味する．

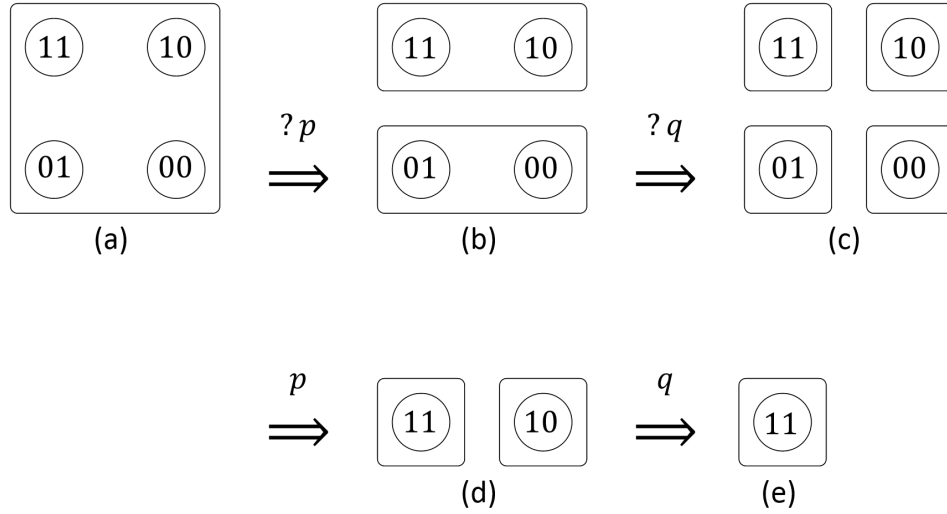


図 4.4 告知による Σ_a の変化

いま Yes/No 疑問文 $?p$ という告知がなされた．このとき Σ_a は変化し， Σ_a 中にあるすべての部分集合に対して p か q が成り立たないといけなないので，図 (b) のように p が成り立つ集合と q が成り立つ集合とに分かれる．しかしエージェント a の information state $\sigma_a = \bigcup \Sigma_a$ は変化しておらず，情報としては何も入っていないことになる．

次に Yes/No 疑問文 $?p$ という告知がなされると， Σ_a はさらに細分化され，図 (c) になる．このときもまた σ_a は変化せず情報は増えないが， p が真であるかどうか， q が真であるかどうかについて明確に issue は区別されている．

いま平叙文 p という告知が行われると，図 (d) になる． p という情報により，エージェント a の information state σ_a は p が成り立たない可能世界を削るように変更される．また， p かどうかという issue は解明される．

最後に平叙文 q という告知が行われ，図 (e) になる．告知により情報が増え， σ_a は q が成り立たない世界を削るように変更され， q かどうかという issue は解明されたといえる．

第5章 IDEL[~]の提案とその実装と解析

本章では、まず公開告知論理の構文論，クリプキ意味論を実装したシステムを説明し，テストケースとして“誕生日の謎”を解析し，可能世界の動的変化が実装により視覚化できていることを確認する．IDEL とそれを改良した体系 IDEL[~] を新たに提案し，さらに推理小説の例として挙げた『不連続殺人事件』の単純化した問題に対して形式化して解析していく．

5.1 実装

公開告知論理の構文論，クリプキ意味論を実装したシステムについて述べる．このシステムはシステム内にクリプキモデルをあらかじめ構築しておき，論理式を入力することにより，論理式の真偽とクリプキモデルの画像を出力する．システムは `haskell` と `Graphviz` で実装し，入力された論理式の解析とその論理式により変化したクリプキモデルの構築に `haskell`，モデル画像の生成に `Graphviz` を使用している．実装環境は以下の通りである．

- OS : Windows7 Enterprise 64bit
- CPU : Intel(R) Core(TM)2 Quad CPU 3.00GHz
- メモリ : 16GB

またシステムの具体的な処理の流れは以下のようになる．

1. 論理式解析
2. 論理式計算
3. クリプキモデルの更新
4. モデル画像の出力

5.1.1 システムの構成

本システムは `haskell` を中心に構成し，画像の出力のみ `Graphviz` を用いている．モデルケースとして第三章に挙げた誕生日のパズルを実装したプログラムを用いて解きながら，

そのシステムについて説明していく．まずはクリプキモデルの設定をプログラム内部で行う．

- エージェントの集合 $A = \{a, b\}$
- 可能世界の集合 $W = \{3/4, 3/5, 3/8, 6/4, 6/7, 9/1, 9/5, 12/1, 12/2, 12/8\}$
- A の到達可能性関係 $R_a = \{((x, y), (x', y')) \in W \times W \mid x = x'\}$
- B の到達可能性関係 $R_b = \{((x, y), (x', y')) \in W \times W \mid y = y'\}$
- 付値 $V(x/y \text{ is the } C\text{'s birthday}) = \{(x, y)\}$

このモデルを計算機上で実装すると、図 5.1 になる．

エージェント、可能世界、到達可能性関係のそれぞれの情報を“birthAgent”, “birthWorld”, “birthRelation” が持っている．可能世界は対で表現していて、3 月 4 日という可能世界は (3,4) として表し、到達可能性関係は (エージェント, 可能世界, 可能世界) の形を持っていて、2 項目の可能世界から 3 項目の可能世界に対し繋がっている．また一つ一つの命題に対しての付値が “valueP (p)” として与えられ、可能世界を要素に持つ集合として定義され、その valueP (p) を集めた集合を “birthValues” として与える．このモデルの可変な情報は “Frame” に格納されており、Frame には可能世界と、到達可能性関係についての情報を保持している．いま設定したモデルの情報を取り出すと、実行画面上で図 5.2 のように出力される．

この情報を視覚的に捉えるために、graphviz によって画像を出力する．Frame の情報を用いて表示したい可能世界の情報と到達可能性関係の情報を使って、内部でプログラミングしたパーサーにより graphviz で読み込める txt ファイルを作り、その txt ファイルを実行し画像を出力する．設定した Frame の情報を用いてできた画像が図 5.4 である．

出来上がったモデルの付値、可能世界に対し告知を与えると真かどうかを評価する．例えば、いま可能世界 (12,2) にいるとする．このとき日にちを知っている B は到達可能性関係により他の可能世界と繋がっていないので、瞬時に日付が分かる．告知を「C の誕生日が 12 月 2 日だと分かる」とすると、プログラムにより、論理演算を行い真か偽であるかを返してくれる．図 5.4 はその実行画面である．

このようにして “valueF” 関数を用いて真偽値を評価する．

公開告知による動的変化

動的な変化を実装するにあたり必要なのはその告知の真理集合である．Frame の中にあるすべての可能世界に対し告知による真偽値を計算し、真になる可能世界だけを集めたものが真理集合となり、その集合を更新されたモデルの可能世界とする．さらに Frame の到達可能性関係でいま偽になる可能世界が入っているものを取り除いてできる集合が更新

```

41 ----- valuation & frame of birthday -----
42
43 birthAgent = ["month","day"]
44 birthWorld = [(3,4),(3,5),(3,8),(6,4),(6,7),(9,1),(9,5),(12,1),(12,2),(12,8)]
45
46 birthRelation = [("month",x,y)|x <- birthWorld, y <- birthWorld, fst x == fst y]++
47                [("day",x,y)|x <- birthWorld, y <- birthWorld, snd x == snd y ]
48
49
50 valueP :: Formula -> WorldList
51 valueP (Atom "m3d4") = [(3,4)]
52 valueP (Atom "m3d5") = [(3,5)]
53 valueP (Atom "m3d8") = [(3,8)]
54 valueP (Atom "m6d4") = [(6,4)]
55 valueP (Atom "m6d7") = [(6,7)]
56 valueP (Atom "m9d1") = [(9,1)]
57 valueP (Atom "m9d5") = [(9,5)]
58 valueP (Atom "m12d1") = [(12,1)]
59 valueP (Atom "m12d2") = [(12,2)]
60 valueP (Atom "m12d8") = [(12,8)]
61
62 birthFrame = Frame birthWorld birthRelation
63
64 birthValues = [valueP (Atom "m3d4"),
65                valueP (Atom "m3d5"),
66                valueP (Atom "m3d8"),
67                valueP (Atom "m6d4"),
68                valueP (Atom "m6d7"),
69                valueP (Atom "m9d1"),
70                valueP (Atom "m9d5"),
71                valueP (Atom "m12d1"),
72                valueP (Atom "m12d2"),
73                valueP (Atom "m12d8")]
74

```

図 5.1 モデルの設定

```

*Main> birthFrame
Frame {world = [(3,4),(3,5),(3,8),(6,4),(6,7),(9,1),(9,5),(12,1),(12,2),(12,8)], relation = [(~month~, (3,
4),(3,4)), (~month~, (3,4),(3,5)), (~month~, (3,4),(3,8)), (~month~, (3,5),(3,4)), (~month~, (3,5),(3,5)), (~mo
nth~, (3,5),(3,8)), (~month~, (3,8),(3,4)), (~month~, (3,8),(3,5)), (~month~, (3,8),(3,8)), (~month~, (6,4),(6,4
)), (~month~, (6,4),(6,7)), (~month~, (6,7),(6,4)), (~month~, (6,7),(6,7)), (~month~, (9,1),(9,1)), (~month~, (
9,1),(9,5)), (~month~, (9,5),(9,1)), (~month~, (9,5),(9,5)), (~month~, (12,1),(12,1)), (~month~, (12,1),(12,2)
), (~month~, (12,1),(12,8)), (~month~, (12,2),(12,1)), (~month~, (12,2),(12,2)), (~month~, (12,2),(12,8)), (~
month~, (12,8),(12,1)), (~month~, (12,8),(12,2)), (~month~, (12,8),(12,8)), (~day~, (3,4),(3,4)), (~day~, (3,4
),(6,4)), (~day~, (3,5),(3,5)), (~day~, (3,5),(9,5)), (~day~, (3,8),(3,8)), (~day~, (3,8),(12,8)), (~day~, (6,4),(3,4
)), (~day~, (6,4),(6,4)), (~day~, (6,7),(6,7)), (~day~, (9,1),(9,1)), (~day~, (9,1),(12,1)), (~day~, (9,5),(3,5)), (~
day~, (9,5),(9,5)), (~day~, (12,1),(9,1)), (~day~, (12,1),(12,1)), (~day~, (12,2),(12,2)), (~day~, (12,8),(3,8)), (~
day~, (12,8),(12,8))]}

```

図 5.2 Frame

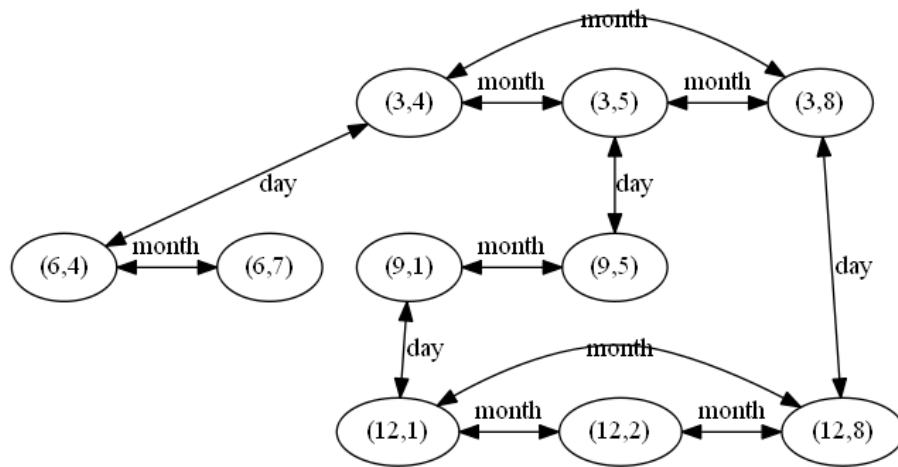


図 5.3 初期状態

```
*Main> valueF birthFrame birthValues (12,2) (Box "day" (Atom "m12d2"))
True
```

図 5.4 告知の真偽値

```

98
99 ----- pが真になる世界, 偽になる世界 -----
100 trueWorld :: Frame -> Valuation -> Formula -> WorldList
101 trueWorld f v p = [x | x <- (world f), (valueF f v x p) == True]
102
103 falseWorld :: Frame -> Valuation -> Formula -> WorldList
104 falseWorld f v p = [x | x <- (world f), (valueF f v x p) == False]
105
106 ----- pが真になる到達可能性関係 -----
107 trueRelation :: Frame -> Valuation -> Formula -> Relation
108 trueRelation f v p = difference (relation f)
109   (union [(ag,u,w) | (ag,u,w) <- (relation f), (w `elem` (falseWorld f v p))],
110    [(ag,w,u) | (ag,w,u) <- (relation f), (w `elem` (falseWorld f v p))])
111
112
113 ----- modifying frame -----
114
115 modifyFrame :: Frame -> Formula -> Frame
116 modifyFrame f p = Frame (trueWorld f birthValues p) (trueRelation f birthValues p)
117
```

図 5.5 モデルの更新

されたモデルでの到達可能性関係となる．告知により変化したモデルの Frame には更新された可能世界の集合と到達可能性関係を持つ．

図 5.5 は告知の真理集合と真な到達可能性関係，それによって更新されたモデルを作る関数 “modifyFrame” について記述したものである．この関数に Frame と告知を入れることで，新しいモデルの Frame が生成される．

5.1.2 誕生日の謎をシステムを用いての検証

初期状態は図 5.4 であり，この図から告知によりどのように変化するか検証していく．まず告知をプログラム上で表現できる形に変換していく．第三章において形式化された告知を告知演算子 $[]$ を省いて表記すると，

- 「エージェント $z \in \mathcal{A}$ は C の誕生日を知っている」

$$\bigvee_{(x,y) \in [cob]} K_z(x/y \text{ is the } C\text{'s birthday})$$
- 1. A 「B は C の誕生日がわからないはずです」

$$\text{ann1} = K_a(\bigwedge_{(x,y) \in [cob]} \neg K_b(x/y \text{ is the } C\text{'s birthday}))$$
- 2. B 「今の A さんの言葉 C の誕生日がわかりました」

$$\text{ann2} = \bigvee_{(x,y) \in [cob]} K_b(x/y \text{ is the } C\text{'s birthday})$$
- 3. A 「私も C の誕生日がわかりました」

$$\text{ann3} = \bigvee_{(x,y) \in [cob]} K_a(x/y \text{ is the } C\text{'s birthday})$$

になる．

これをプログラム上で表現したものが図 5.6 になる．

「エージェント $z \in \mathcal{A}$ は C の誕生日を知っている」は，命題の集合を “meidai” として表し，その命題に対して認識演算子を用い，さらに出来上がった論理式の論理和をとることで表現され，“kmm1”，“kdd1” のように記述できる．また ann1, ann2, ann3 はそれぞれ “announce1”，“announce2”，“announce3” として記述している．

初期状態のモデルに対し ann1 を与え更新されたモデルが “modifyFrame1” になる．さらに “modifyFrame1” に対し ann2 を行ったものが “modifyFrame2”，そのモデルに対し ann3 を行ったのが “modifyFrame3” になる．

これらの更新されたモデルを図で表現する．

この出来上がった図 5.7, 5.8, 5.9 は第三章において解いていく過程で用いた図と同じものとなっており，このシステムは正しく実装できている．このようにして，意味論を実装し，可能世界や到達可能性関係や付値を持ったクリプキモデルを与え，告知を形式化しその真偽値を計算することにより，動的に計算機によりシミュレートすることができたといえる．

```

158 meidai = [(Atom "m3d4"), (Atom "m3d5"), (Atom "m3d8"), (Atom "m6d4"), (Atom "m6d7"), (
159   Atom "m9d1"), (Atom "m9d5"), (Atom "m12d1"), (Atom "m12d2"), (Atom "m12d8")]
160 kmm1 = Orlist [Box "month" y|y <- meidai]
161 kdd1 = Orlist [Box "day" y|y <- meidai]
162
163 announce0 = Not kdd1
164 announce1 = Box "month" (announce0)
165 announce2 = kdd1
166 announce3 = kmm1
167
168 modifyFrame0 = modifyFrame birthFrame announce0
169 modifyFrame1 = modifyFrame birthFrame announce1
170 modifyFrame2 = modifyFrame modifyFrame1 announce2
171 modifyFrame3 = modifyFrame modifyFrame2 announce3
172

```

図 5.6 告知と変化した Frame

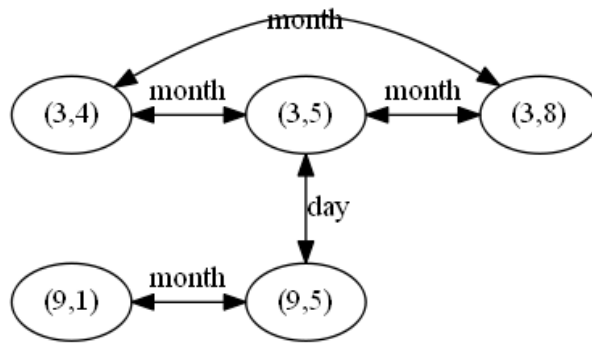


図 5.7 modifyFrame1

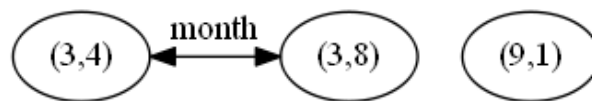


図 5.8 modifyFrame2

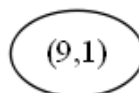


図 5.9 modifyFrame3

5.2 推理小説の解析

二章にて挙げた『不連続殺人事件』を簡単化した問題をもう一度載せる。

容疑者は A, B, C, D, E

ある村で殺人事件が 3 つ起こる

一つ目の殺人のとき, A, B は村の外に出ており, 目撃者にも確認が取れていて確固としたアリバイがある. D と E は共に一緒にいたと証言. C のみ部屋におりアリバイがない.

二つ目の殺人のとき, 怪しいとされる C, D, E はそれぞれ自分の部屋にいて, B がそれを歩きながら廊下で監視していた. B は 3 人とも部屋を出ることはなかったと証言. C, D, E もまた B が廊下を離れることはなかったと証言. A のみアリバイがない.

三つ目の殺人のとき, A, C, D は一緒にいたと互いに証言した. B と E にはアリバイがなかった.

犯人は一人もしくは, その共犯が一人だけいる可能性がある. さあ犯人は誰。

5.2.1 IDEL を用いた可能性の削減

IDEL を用いて可能性を issue と可能世界という両方の尺度から削減できるということを示してきた. ここでは issue の変化が問題に対してどのように使うことができるかを示していく.

形式化

一可能世界に対し一命題にする. つまり, w_a という可能世界では「A が犯人である」という命題 h_a が成り立っているモデルを考える. また容疑者として A, B, C, D, E の 5 人がいるが, 犯人には犯人の到達可能性関係があり, この時点ではまだ犯人は分かっていないので, 容疑者に対して到達可能性関係は与えられない. よって到達可能性関係は読者 r のもののみを設定する. また到達可能性関係 $\Sigma_r(w)$ はすべての可能世界で同じであり, Σ_r として表す.

- 可能世界 w の集合 $W = \{w_a, w_b, w_c, w_d, w_e\}$
- issue $\Sigma_r = \wp(W)$
- 付値 $V(h_x) = \{w_x\}$

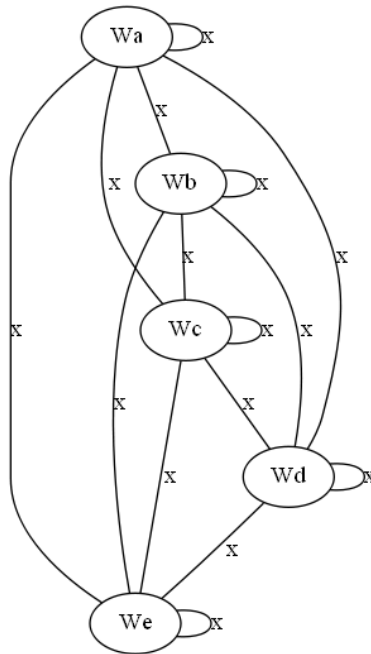


図 5.10 各可能世界とその issue

```
*Main> issues cFrame
[[["We"],["Wd"],["Wd","We"],["Wc"],["Wc","We"],["Wc","Wd"],["Wc","Wd","We"],["Wb"],["Wb",
"We"],["Wb","Wd"],["Wb","Wd","We"],["Wb","Wc"],["Wb","Wc","We"],["Wb","Wc","Wd"],["Wb",
"Wc","Wd","We"],["Wa"],["Wa","We"],["Wa","Wd"],["Wa","Wd","We"],["Wa","Wc"],["Wa","Wc",
"We"],["Wa","Wc","Wd"],["Wa","Wc","Wd","We"],["Wa","Wb"],["Wa","Wb","We"],["Wa","Wb",
"Wd"],["Wa","Wb","Wd","We"],["Wa","Wb","Wc"],["Wa","Wb","Wc","We"],["Wa","Wb","Wc","Wd",
"],["Wa","Wb","Wc","Wd","We"]]]
*Main> length it
31
```

図 5.11 issue の要素

図 5.10 は設定したモデルの可能世界とその issue を表したものである．issue は本来グループになるが，作図の都合上同じ issue に入っている可能世界同士を線で結んでいる．issue の最大の要素は W になっているので，すべての可能世界同士が issue によって繋がっている．図 5.11 はその issue の中身とその要素数を表示させたものである．

いま読者の issue はすべての可能世界のべき集合であり図 5.11 よりその要素数は 31 個であることが分かる．この issue はある命題が成り立つ集合であるが，このモデルでの意味づけを変え，その issue に含まれる要素一つ一つを共犯の組にする．つまり issue の中に $\{w_a, w_b\}$ があったとすると，その可能世界の集合は「A, B が共犯である」という意味を持つものとする．問題の「犯人は一人もしくは，その共犯が一人だけいる可能性がある」を形式化する．犯人は二人以下であるので，downward closure の性質が発揮でき，「犯人は二人以下の組み合わせであり，犯行を行ったのが組み合わせの中のどちらかだ」と考えると， $?$ を用いて次のように表せる．

$$\text{ann1} = \{x \vee y \mid x, y \in \{h_a, h_b, h_c, h_d, h_e\}\}$$

この告知を行ったとき，可能世界の数は変化しないが Σ_r が告知が成り立つように $[\text{ann1}]$ により制限される．

```
*Main> issues modifyFrame1
[[["We~"],["Wd~"],["Wd~","We~"],["Wc~"],["Wc~","We~"],["Wc~","Wd~"],["Wb~"],["Wb~","We~"],["Wb~","Wd~"],
["Wb~","Wc~"],["Wa~"],["Wa~","We~"],["Wa~","Wd~"],["Wa~","Wc~"],["Wa~","Wb~"]]
*Main> length it
15
```

図 5.12 ann1 後の Σ_r

図 5.12 で表示された Σ_r をみると可能世界の長さが 3 以上ある可能世界の集合が消えて，長さ 2 以下，つまり犯人は二人以下の組み合わせが残る．

この作られた issue の要素の一つ一つに命題を与えていけば PAL で解くこともできるが，それでは issue を用いて犯人の組み合わせを減らすまでもなく，5 人いる中から 2 人以下であるという意味で ${}_5C_2 + {}_5C_1$ の動作を行ってクリプキモデルを作ってしまう issue を使うまでもなくできてしまう．それではより複雑な問題には対処できない．よって issue という可能世界の集合を用いて可能性を削減していき特定の集合のみが真となる場合を推論できるようしなければならない．そのためには issue の要件である downward closure を失くす必要がある．downward closure は数学的にも理にかなってはいるが，推理小説を解くために issue を共犯の組の意として捉えると制限が強すぎるのである．例えば，このモデルにおいて犯人が A と B だとすると downward closure の性質により，A の単元集合と B の単元集合も含まれてしまう．A と B が共犯でやっと思える犯罪については推論することができないのである．また，問題設定が犯人は二人以下であるからこそ downward closure の性質を利用できたが，犯人は三人以上の場合にはこれは用いることはできない．

これでは普遍性がある推論とはいえない．次のパートで downward closure の性質を破る論理について提案する．

5.2.2 IDEL[~]

IDEL に新たに論理式 φ を不成立にする古典的否定演算子 “ $\sim$ ” を加える．意味論で表すと，

$$M, s \models \sim \varphi \iff M, s \not\models \varphi$$

この演算子を入れることにより、「ある可能世界で φ がなりたっている」ということを表現できる．今まで issue の持っていた downward closure の性質は，基本となる論理式 $M, s \models \varphi$ の「すべての可能世界で φ がなりたつ」の意味によりその性質を保証されていたが，「ある可能世界で φ がなりたっている」が表現できるため，issue の中である可能世界の集合では成り立たないが他の可能世界を含めた集合を考えると成り立つ場合が考えられ，issue は downward closure ではなくなる．よってこの issue の代わりとなる可能世界の集合の集合に対し新たに擬似という意味を頭につけ，p-issue (pseudo issue) と呼ぶことにする．また p-issue に基づいた IDEL を IDEL[~] と名づける．

構文論

issue の変化に伴い構文論も書き換える必要がある．平叙文全体の集合を $\mathcal{L}_!$ ，疑問文全体の集合を $\mathcal{L}_?$ とし，疑問文 $\alpha, \beta, \gamma \in \mathcal{L}_!$ と平叙文 $\mu, \nu, \lambda \in \mathcal{L}_?$ と限定しない文 $\varphi, \psi, \chi \in \mathcal{L}_! \cup \mathcal{L}_?$ を定義すると，以下に示す構文となる．

1. すべての $p \in P$ に対し $p \in \mathcal{L}_!$
2. $\perp \in \mathcal{L}_!$
3. $\alpha_1, \dots, \alpha_n \in \mathcal{L}_!$ ならば $?\{\alpha_1, \dots, \alpha_n\} \in \mathcal{L}_?$
4. $\varphi \in \mathcal{L}_\circ$ かつ $\psi \in \mathcal{L}_\circ$ ならば $\varphi \wedge \psi \in \mathcal{L}_\circ$ ，ここで $\circ \in \{!, ?\}$
5. $\alpha \in \mathcal{L}_!$ かつ $\varphi \in \mathcal{L}_\circ$ ならば $\alpha \rightarrow \varphi \in \mathcal{L}_\circ$ ，ここで $\circ \in \{!, ?\}$
6. $\circ \in \{!, ?\}$ であり， $\varphi \in \mathcal{L}_\circ$ かつ $a \in A$ ならば $B_a \varphi \in \mathcal{L}_!$
7. $\circ \in \{!, ?\}$ であり， $\varphi \in \mathcal{L}_\circ$ かつ $a \in A$ ならば $E_a \varphi \in \mathcal{L}_!$
8. $\circ \in \{!, ?\}$ であり， $\varphi \in \mathcal{L}_\circ$ かつ $\bullet \in \{!, ?\}$ であり， $\psi \in \mathcal{L}_\bullet$ ならば $[\varphi]\psi \in \mathcal{L}_\bullet$
9. $\mathcal{L}_!$ にも $\mathcal{L}_?$ にも属していない文はない
10. $\alpha, \beta \in \mathcal{L}_!$ ならば $\alpha \vee \beta \in \mathcal{L}_!$

11. $\alpha \in \mathcal{L}_!$ ならば $\sim \alpha \in \mathcal{L}_!$

「または」 $\vee$ と「成り立たない」 $\sim$ という演算子でできた論理式を構文に加える.

意味論

クリプキモデルの到達可能性関係が少し意味が変わってくる. IDEL では $(\Sigma_a)_{a \in A}$ は可能世界の集合 W から Π への関数であったが, ここでは W から $\wp(\wp(W))$ への写像とする. また到達可能性関係の Factivity を落とし, 認識演算子 K の代わりに, 信念演算子 B を用いる. クリプキモデルを $M = (W, (\Sigma_a)_{a \in A}, V)$ とすると到達可能性関係は,

- $(\Sigma_a)_{a \in A} : W \rightarrow \wp(\wp(W))$ となる写像であり, ここで可能世界 w からの写像である $\Sigma_a(w)$ には以下の性質がある.

– Introspection : すべての $w, v \in W$ に対し, $v \in \sigma_a(w)$ ならば $\Sigma_a(v) = \Sigma_a(w)$

$\sigma_a(w) := \bigcup \Sigma_a(w)$ はエージェント a の w における information state を表している

解明条件は以下のように書き換えられる.

1. $M, s \models p \iff$ すべての可能世界 $w \in s$ に対して $w \in V(p)$
2. $M, s \models \perp \iff s = \emptyset$
3. $M, s \models ?\{\alpha_1, \dots, \alpha_n\} \iff$ ある $1 \leq i \leq n$ に対して $M, s \models \alpha_i$
4. $M, s \models \varphi \wedge \psi \iff M, s \models \varphi$ かつ $M, s \models \psi$
5. $M, s \models \alpha \rightarrow \varphi \iff$ すべての $t \subseteq s$ に対し, $M, t \models \alpha$ ならば $M, t \models \varphi$
6. $M, s \models B_a \varphi \iff$ すべての $w \in s$ に対し $M, \sigma_a(w) \models \varphi$
7. $M, s \models E_a \varphi \iff$ すべての $w \in s$ かつ, すべての $t \in \Sigma_a(w)$ に対して $M, t \models \varphi$
8. $M, s \models [\varphi] \psi \iff M, s \models \varphi$ ならば $M^\varphi, s \models \psi$
9. $M, s \models \alpha \vee \beta \iff$ ある t_1, t_2 があり, $t_1 \cup t_2 = s$ かつ $t_1 \models \alpha$ かつ $t_2 \models \beta$
10. $M, s \models \sim \varphi \iff M, s \not\models \varphi$

ここで $M^\varphi = (W^\varphi, (\Sigma_a)_{a \in A}^\varphi, V^\varphi)$ は以下のように定義される.

- $W^\varphi = W \cap \bigcup [\![\varphi]\!]_M$
- $(\Sigma_a)_{a \in A}^\varphi = (\Sigma_a)_{a \in A} \cap [\![\varphi]\!]_M$

- $V^\varphi(p) = V(p) \cap \bigcup \llbracket \varphi \rrbracket_M$

また Yes/No 疑問文は,

- $M, s \models ?\alpha \iff M, s \models \alpha \text{ または } M, s \models \neg\alpha$

として定義する.

このとき以前の Fact1: $M, s \models \varphi$ かつ $t \subseteq s$ ならば $M, t \models \varphi$, は $\sim \varphi$ を入れたため成り立たない. IDEL においては downward closure の性質により, ある state で成り立つ論理式はその部分集合でも成り立っていたが, 成り立たない論理式について記述できるため, ある state で成り立たない論理式はその部分集合でも成り立たないということがいえない.

また, ある p が成り立つ可能世界 w_1 と p が成り立たない可能世界 w_2 の集合を考えたとき, w_1 において $\sim p$ は成り立たないので $\text{proposition}[\llbracket \sim p \rrbracket M]$ は $\{w_1\}$ を含んでいないが, w_1 と w_2 の集合をみると $\sim p$ が成り立ち, $\llbracket \sim p \rrbracket$ に $\{w_1, w_2\}$ が含まれていて, 図 5.13 のような可能世界の集合の集合ができる.

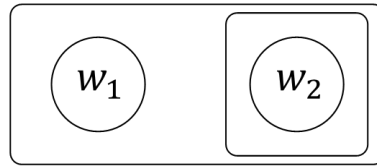


図 5.13 $\text{proposition}[\llbracket \sim \varphi \rrbracket]$

$$\llbracket \sim p \rrbracket = \{\{w_2\}, \{w_1, w_2\}\}$$

このとき $\{w_1, w_2\}$ を含むのに対し, $\{w_1\}$ を含んでいないことから $\llbracket \sim p \rrbracket$ は downward closure の性質を持っていないことが分かる.

そして可能世界に対する付値の関係は以下のことがいえる.

$$M, w \models \neg\varphi \iff M, w \models \sim\varphi$$

演算子 $\sim$ の性質

$\sim$ の意味論を加えることにより, 「state s の中のある可能世界で p が成り立つ」を $\neg$ と $\sim$ を用いて表現することが可能になる. いま p-issue の要素の一つ一つが可能性の選択肢の意味を持っているため, 「state s の中のある可能世界で p が成り立つ」は「 p の可能性がある」と解釈できる.

$$\begin{aligned} M, s \models \sim \neg p &\iff M, s \not\models \neg p \\ &\iff \text{ある集合 } t \subseteq s \text{ に対し, } M, t \models p \\ &\iff \text{ある可能世界 } w \in s \text{ に対し, } w \in V(p) \end{aligned}$$

この性質がいかに p-issue に影響を及ぼすのかを例を用いて説明する．以下のように設定したクリプキモデルを与える．

- 可能世界 w の集合 $W = \{w_1, w_2, w_3\}$
- p-issue $\Sigma_a = \wp(W)$
- 付値 $V(p) = \{w_1, w_2\}$ $V(q) = \{w_1, w_3\}$

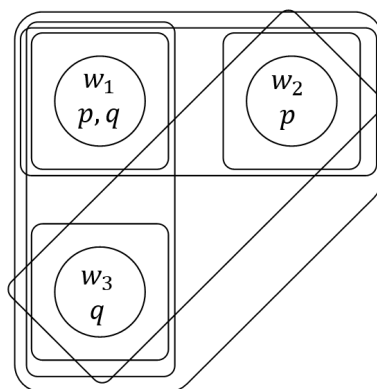


図 5.14 初期状態の p-issue

図 5.14 は設定したクリプキモデルのエージェント a の p-issue である．いま Σ_a は設定により可能世界全体のべき集合であるため downward closure の性質を保持している．このモデルに対し、「 p の可能性がある」という意味の告知 $\text{ann1} = \sim \neg p$ を行う．

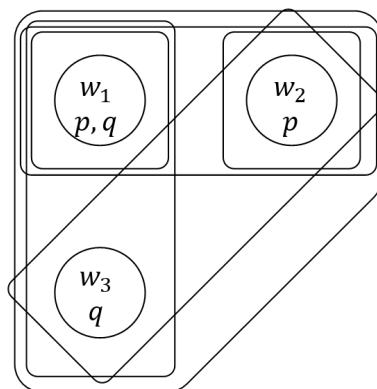


図 5.15 ann1 後の p-issue

図 5.15 は「 p の可能性がある」という告知により変化した p-issue である． p が state s のどの可能世界でもなりたっていない可能世界の集合 $\{w_3\}$ が消える．また $\{w_1, w_3\}$ では,

可能世界 w_1 において p が成り立っているので告知も成り立ち、選択肢として残る。可能世界全体である W も p が成り立つ可能世界を含んでいるため成り立つ。このときの p-issue は $\Sigma_a^{ann1} = \{\{w_1, w_2, w_3\}, \{w_1, w_2\}, \{w_1, w_3\}, \{w_2, w_3\}, \{w_1\}, \{w_2\}\}$ である。可能世界 w_3 は残っているが、 $\{w_3\}$ が無くなり、downward closure の性質がなくなっていることが分かる。

次に「 $\neg p$ の可能性がある」という意味の告知 $ann2 = \sim p$ を行った。

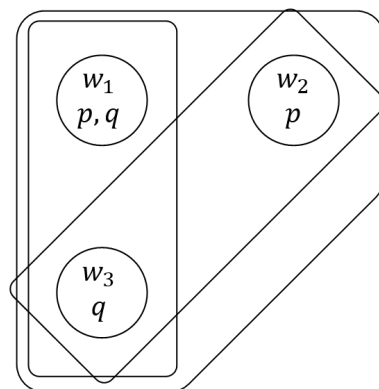


図 5.16 ann2 後の p-issue

図 5.16 は「 $\neg p$ の可能性がある」という告知により変化した p-issue である。 $\neg p$ が成り立っていない可能世界のみの集合 $\{w_1\}, \{w_2\}, \{w_1, w_2\}$ が失われ、p-issue は $\Sigma_a^{ann1ann2} = \{\{w_1, w_2, w_3\}, \{w_1, w_3\}, \{w_2, w_3\}\}$ になる。二つの告知により p-issue の要素は 7 個あったが、徐々に減り 3 個の可能性の選択肢が残された。またこのとき可能世界の単元集合は一つも含んでいない。つまり一つの可能世界だけではなりたたないが、二つの可能世界の集合とすると告知が成り立つということを示している。これは犯人推論において、犯人が複数人おり、一人では犯行不可能だが二人では犯行可能になるという場合に適している。

IDEL~ による解析

IDEL 同様に「犯人は一人またはその共犯がいる」という告知が行われた後から推論を続けていく。いま図 5.12 をみると、issue の要素数は 15 個である。この issue から徐々に可能世界のグループが減っていく様子をシミュレートしていく。

まず告知を検証し形式化する。

一つ目の殺人のとき、A、B は村の外に出ており、目撃者にも確認が取れていて確固としたアリバイがある。D と E は共に一緒にいたと証言。C のみ部屋におりアリバイがない。

この一つの場面を一つの告知により表現する。まず A と B はこの時点ではこの事件において殺人を行っていないことは証明されているが、犯人ではないとは言いきれない。C に

についてはアリバイがないので犯行の可能性はある。また D と E に関しては第二章で推論したとおり、共犯であれば互いに嘘をつくことができるので D と E は共犯の可能性はある。つまりここで犯人の可能性に関して読み取れるのは「C が犯人である可能性はあるか、D と E が共犯である可能性はある」である。ここでまず D と E が共犯の可能性あることを形式化すると、

- 「D と E は共犯の可能性はある」

$$\sim \neg h_d \wedge \sim \neg h_e$$

さらに C についての可能性を含めると、? を用いて表現することができ、

- 「C が犯人である可能性はあるか、D と E は共犯の可能性はある」

$$\text{ann2} = ?\{\sim \neg h_c, \sim \neg h_d \wedge \sim \neg h_e\}$$

二つ目の事件について形式化する。

二つ目の殺人のとき、怪しいとされる C, D, E はそれぞれ自分の部屋にいて、B がそれを歩きながら廊下で監視していた。B は 3 人とも部屋を出ることはなかったと証言。C, D, E もまた B が廊下を離れることはなかったと証言。A のみアリバイがない。

A にはアリバイがなく、C, D, E は B と共犯であれば犯行可能であり、「A は犯人の可能性があり、B と C が共犯、または B と D が共犯、または B と E が共犯である可能性はある」といえる。これを ann2 として表すと、

- 「A は犯人の可能性があり、B と C が共犯、または B と D が共犯、または B と E が共犯である可能性はある」

$$\text{ann3} = ?\{\sim \neg h_a, \sim \neg h_b \wedge \sim \neg h_c, \sim \neg h_b \wedge \sim \neg h_d, \sim \neg h_b \wedge \sim \neg h_e\}$$

三つ目の事件について形式化する。

三つ目の殺人のとき、A, C, D は一緒にいたと互いに証言した。B と E にはアリバイがなかった。

A, C, D が共犯であれば犯行可能であり、B と E にはアリバイがないので、「A と C と D が共犯の可能性があり、B と E は犯人の可能性はある」といえるので、ann4 として形式化すると、

- 「A と C と D が共犯の可能性があり、B と E は犯人の可能性はある」

$$\text{ann4} = ?\{\sim \neg h_a \wedge \sim \neg h_c \wedge \sim \neg h_d, \sim \neg h_b, \sim \neg h_e\}$$

実装したプログラムにより告知によって犯人が導かれるか検証していく。

告知 ann2 を行った後、可能世界の数は変化しないが、p-issue の要素が変化する。

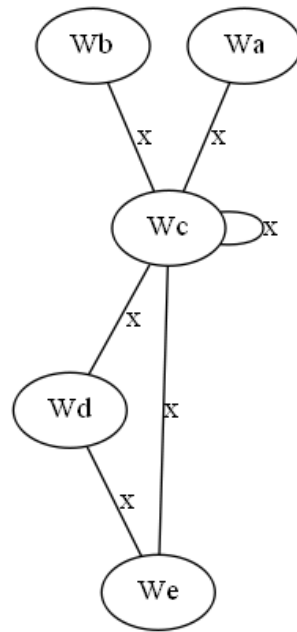


図 5.17 ann2 後の可能世界と p-issue

```

*Main> issues modifyFrame2
[["Wd","We"],["Wc"],["Wc","We"],["Wc","Wd"],["Wb","Wc"],["Wa","Wc"]]
*Main> length it
6

```

図 5.18 ann2 後の Σ_r

図 5.17 は告知 ann2 を行った後の可能世界と、p-issue に含まれている可能世界の集合の要素同士を線で結んだものであり、図 5.18 は読者が抱いている p-issue である。発言により「C が犯人」という可能世界 w_c を含んでいる可能世界の集合 $\{w_a, w_c\}, \{w_b, w_c\}, \{w_c, w_d\}, \{w_c, w_e\}, \{w_c\}$ と、「D が犯人である」かつ「E が犯人である」という命題を持っている可能世界の集合 $\{w_d, w_e\}$ が残る。この制限されたモデルに対して告知 ann3 を行う。

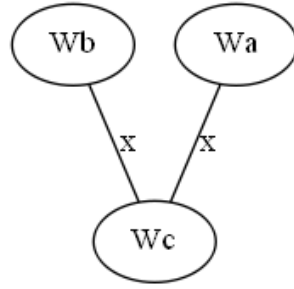


図 5.19 ann3 後の可能世界と p-issue

```

*Main> issues modifyFrame3
[["Wb","Wc"],["Wa","Wc"]]
*Main> length it
2
  
```



図 5.20 ann3 後の Σ_r

図 5.19 が告知 ann3 を行った後の可能世界であり、「D が犯人である」という命題を持っている可能世界 w_d と、「E が犯人である」という命題を持っている可能世界 w_e が削られる。従って D と E は犯人でないということが分かる。また図 5.20 をみると、読者の p-issue の中には $\{w_a, w_c\}, \{w_b, w_c\}$ しか残っておらず、C が単独犯という可能性も消えていることが分かる。このモデルで告知 ann4 を行う。

最後に告知 ann4 を行ってできる可能世界とその p-issue が図 5.21, 図 5.22 になる。可能世界は w_b と w_c だけが残る、また p-issue は $\{w_b, w_c\}$ のみとなり、B と C が共犯である可能性だけが残る、B と C が共犯であったことが分かる。

よって新たに用いた演算子 $\sim$ によって告知を形式化し、その真偽値を計算することにより、動的に計算機によりシミュレートすることができたといえる。

5.3 IDEL $\sim$ の汎用性

IDEL $\sim$ は IDEL の issue の性質である downward closure を落としたものを p-issue とし、issue では情報量はゼロであったが、p-issue には情報量があり、その p-issue の要素が削ら

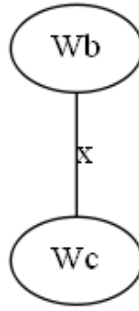


図 5.21 ann4 後の可能世界と p-issue

```

*Main> issues modifyFrame4
[["Wb", "Wc"]]
*Main> length it
1
  
```

図 5.22 ann4 後の Σ_r

れていくことを可能性の削減と捉えることができた．ここでは IDEL[~] の汎用性を示すためにいくつかの例を用い，それを解くことにより説明する．

例えばある殺人事件があったとする．そこでは被疑者は薬を飲まれたうえでナイフにより殺されていた．容疑者が 4 人いる．それぞれ A, B, C, D は属性を持っていて，薬を扱えるのは A と B，ナイフを扱えるのは B と C と D とする．薬を扱えるという命題を m ，ナイフを扱えるという命題を k ，エージェントは犯人であるという命題を $(h_a)_{a \in A}$ とし，設定したクリプキモデルは以下になる．

- 可能世界 w の集合 $W = \{w_a, w_b, w_c, w_d\}$
- 読者 r の p-issue $\Sigma_r = \wp(W)$
- 付値 $V(m) = \{w_a, w_b\}$ $V(k) = \{w_b, w_c, w_d\}$ $V(h_x) = \{w_x\}$

図 5.23 がこのモデルの初期状態である．事件を単純に考えると，被害者は薬を飲まれたうえでナイフで殺されたのでこれは薬を扱える人物であり，かつナイフも扱える人物であると読み取れる．IDEL のみを用いてこの告知「薬を扱え，ナイフも扱える人物である」 $m \wedge k$ をすると，B のみが犯人として浮かび上がる．計算機で実装したものが図??になり，やはり B が犯人であることが確かめられる．

だがここで B には完全なアリバイがあることが分かってしまった．今可能世界は w_b しか残っていないので可能世界がなくなり，この推論は間違っていたことが分かる．

次に IDEL[~] を用いて告知を形式化する．やはり被害者は薬を飲まれたうえでナイフで殺されていたのだが，薬を扱える人物と，ナイフを扱える人物が別々に被疑者を殺害し

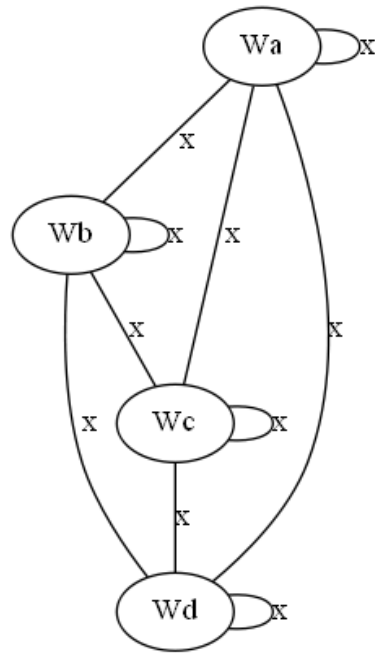


図 5.23 初期状態

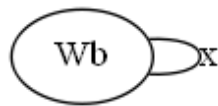


図 5.24 告知 $m \wedge k$ 後の issue

たとも読み取ることができる． よって p-issue の要素を犯行を行ったグループの命題がなりたつ可能世界の集合と捉えることができる． 演算子 $\sim$ を用いて告知を形式化し， その告知「薬を扱える可能性があり， ナイフも扱える可能性がある」 $\sim \neg m \wedge \sim \neg k$ を計算機上で行うと， p-issue の出力画面は図 5.25 になる．

```
*Main> issues modifyFrame2
[[["Wb"],["Wb","Wd"],["Wb","Wc"],["Wb","Wc","Wd"],["Wa","Wd"],["Wa","Wc"],["Wa","Wc","Wd"],
["Wa","Wb"],["Wa","Wb","Wd"],["Wa","Wb","Wc"],["Wa","Wb","Wc","Wd"]]
```

図 5.25 告知 $\sim \neg m \wedge \sim \neg k$ 後の p-issue

このとき p-issue は薬を扱えてナイフも扱える B が犯人である可能世界 w_b が含まれた可能世界の集合と， 薬が扱える A とナイフが扱える B， C， D のいずれかが犯人である可能世界を含んだ集合とで構成される．

また B には完全なアリバイがあることが分かったため， 告知 $\neg h_b$ を行うと， 図 5.26 となる．

```
*Main> issues modifyFrame3
[[["Wa","Wd"],["Wa","Wc"],["Wa","Wc","Wd"]]
```

図 5.26 告知 $\neg h_b$ 後の p-issue

三つの可能性の選択肢が残ることになる． 少なくとも A は犯人であり被疑者に薬を盛っていたことが分かり， その後で C か E のどちらか， もしくは両方がナイフにより被疑者を殺害したことがこの結果より予想される．

このようにして， IDEL を用いただけでは可能世界がなくなり推論が止まってしまうモデルにおいても， IDEL $\sim$ を用いることで共犯の可能性を考慮し， 推論を進め可能性の選択肢を残すことが可能となる．

第6章 終わりに

6.1 考察

本研究では、推理小説の解析というテーマのもと、読み手の知識がどのように変化していくのかという観点で推論を形式化するために、まず公開告知論理により“誕生日の謎”を例にとり、エージェントの知識が告知により動的に変化していく様子を形式化した。更に inquisitive semantics の考え方を取り入れ、疑問文に対して意味づけを行える論理である IDEL について述べ、その“inquisitive”という概念を捉えるために用いられる issue に対し、問題解決に必要であったため downward closure の制限を取り壊し得られた論理 IDEL $\sim$ を提案した。この IDEL $\sim$ を用いて坂口安吾の推理小説『不連続殺人事件』におけるメイントリックに対し、事件の一つ一つを告知として捉えることで選択肢を制限していくという統一的な解法を与えることができた。また告知を形式化した論理式を用いて計算機上で haskell により論理システムの実装を行い、新しく取り入れた演算子 $\sim$ の動作を確認した。システムでは入力した論理式に対しモデルがどのように変化していくかをクリプキモデルの画像を出力させることで表現した。結果として単純化したものであるが推理小説を計算機上で解析することができ、告知が知識状態に対してどのように変化しているのかを確認することができた。よって基礎的数学的関心に閉じていた動的信念論理を、推理小説の解析という現実の人工知能の問題に応用することができたといえる。

6.2 今後の課題

今後の課題について以下の問題が残っている。

- 本研究で取り入れた論理である IDEL $\sim$ について汎用性を示したものの、その性質は捉え切れなかった部分が多くあると思われる。また issue の性質である downward closure の性質を取り除いたことで出てくる弊害というものも十分には記述することができていないため、それらについて詳しく調査する必要がある。
- 今回推理小説の解析として『不連続殺人事件』を単純化した問題に対して行ったが、推理小説全体を構造解析できるとは言えない。また実用性を考えれば文章中の有意義な情報を小説の中の自然言語から取り出し、その自然言語に対して構文解析を行い論理式を生成することが必要となってくる。小説全体を形式化するために、自然言語からの推論システムを組み上げることが今後の課題である。

謝辞

本論文を作成するにあたり，熱心なご指導を頂いた東条敏教授，佐野勝彦助教に深く感謝いたします。haskell の手解きをしていただきました先輩である野村氏，そして多くのご指摘を下さいました東条研究室の先輩・同期の皆様に感謝いたします。

参考文献

- [1] Hans van Ditmarsch, Wiebe van der Hoek, Barteld Kooi, *Dynamic Epistemic Logic*, Synthese Library, Vol.337.
- [2] 東条敏, 『言語・知識・信念の論理』, 人工知能学会, 2005.
- [3] Ivano Ciardelli, Jeroen Groenendijk, and Floris Roelofsen, Inquisitive semantics, NASSLLI 2012 Lecture Notes, June 25, 2012.
- [4] H. Freudenthal, Formulation of the sum-and-product problem, *Nieuw Archief voor Wiskunde*, 3(17): 152, 1969
- [5] H.P. van Ditmarsch, J.Ruan, R.Verbrugge, Sum and Product in Dynamic Epistemic Logic, *Journal of Logic and Computation*, Volume 18, Issue 4, pp 563-588.
- [6] アーサー・コナン・ドイル, 『シャーロックホームズの事件簿』, 白面の兵士, 新潮文庫, 延原謙訳, 1953.
- [7] 坂口安吾 『不連続殺人事件』 角川書店, 1974.
- [8] Ivano Ciardelli, and Floris Roelofsen, Inquisitive Dynamic Epistemic Logic, *Synthese*, March, 2014.
- [9] Johan van Benthem, and Stefan Minica, Toward a Dynamic Logic of Questions, *Journal of Philosophical Logic*, Volume 41, Issue 4, pp 633-669