OCC theory に基づくエージェントの感情表現と時間経過に関する論理モデル

奈良女子大学大学院　人間文化研究科　情報科学専攻
新出研究室　16840021　今井　那緒

平成 30 年 1 月 31 日

目 次

概 要

近年、エージェントが人間とコミュニケーションを取ることが重要視されており、その一環としてエージェントに感情を持たせることが求められている。OCC theory と呼ばれる、心理学的見地に基づいた人間の感情を 22 種類にモデル化した理論と、この理論で扱われている感情を論理式として形式化し、さらにこれを基に感情の強さを取り入れた研究がある。しかし、これらの研究では時間経過による感情の強さの減衰を扱っていないこと、また、OCC theory で分類された 22 種の感情のうち love と hate の 2 種類は扱えなかったことから、人間らしい感情表現としては不十分であった。そこで、本論文では、従来研究を基に論理式として形式化されていなかった 2 種の感情の形式化と、時間経過による感情の強さの減衰を表現できる論理体系を提案し、この体系で表現される感情の性質などについて述べる。

1 はじめに

近年、人間らしい感情を持つロボットの開発が進んでいる。現実世界において、ロボットが人間とコミュニケーションをとるためには、ロボットが感情を持っていることが望ましい。ロボットが人間らしい感情を持つことで、自ら人間に近い行動を決定することが求められる。

周囲の環境が常に変化する現実世界において、問題解決のために行動するロボットの実現として BDI アーキテクチャーが有効である [6]。BDI アーキテクチャーとは、人間行動を信念、願望、意図の 3 つの心的状態でモデル化した BDI モデル [5] による行為決定方式を計算機上で実現したものである。それにより、人間の合理的思考に基づいた行動をとる BDI エージェントを構築することが可能である。また、この BDI モデルは、BDI logic という論理モデルを持つ。

一方、OCC theory[4] と呼ばれる、人間の感情を分類する理論がある。この OCC theory では、信念や願望などの心的状態を用いて人間の感情を分類、特徴付けており、論理モデルによって表現が可能で、BDI logic を持つ BDI モデルとの親和性が良い。

Adam らの研究 [1] では、OCC theory の理論を BDI モデルに取り込むことで、感情を論理式でモデル化している。この研究で、BDI logic に「不確実だが起こると期待されている事象」等の新たなオペレータを複数導入することで、OCC theory で扱われている 22 種類のうち 20 種の感情を論理式として形式化し、BDI モデルに取り込んでいる。そこで、我々は、Adam らの形式化を踏まえ、BDI アーキテクチャに基づく記述言語 AgentSpeak の処理系である Jason[3] で実現、ライブラリを実装した [7]。この [7] では、1 つの行動につき 1 つの感情しか生起できず、また、複合的な感情については扱っていなかった。それに加えて、生起する感情には度合いがなかった。そのため、[9] では、度合い付きの感情が生起できるような論理体系の提案をし、[8] で、その論理体系に基づいた実装を行った。さらに、[10][11] では、その度合いが時間経過によって減衰し、度合いが 0 になった感情の削除を実装した。しかしこれらの研究 [7][8][9][10][11] では、時間経過による感情の強さに対する論理体系が提案されておらず、また、人間の持つ主要な感情である好意や嫌悪についての感情 (love, hate) が表現できず、人間らしい感情表現としては不十分である。加えて、その他の感情についても、OCC theory では本来、好意と嫌悪の感情を生起条件とする感情があるが、従来研究では表現できなかったため、その生起条件を省略して形式化を行っていた。

そこで本論文では、Adam らの形式化を踏まえつつ、OCC theory で定義されている感情のうち形式化されていなかった 2 種の感情の形式化と、その感情を生起条件とする感情の再形式化、時間経過による感情の強さの減衰を実現する論理体系を提案する。また、本研究は同研究室、生活環境学部情報衣環境学科生活情報通信科学コース 4 年生、浅井 [12] と塚本 [13] との共同研究である。[12][13] では、本論文で述べる論理体系を基に、Jason を用い新たに形式化、再形式化した感情の実装、妥当性の検証を行っている。

2 OCC theory

本章では、先行研究 [1] と [2] でも用いられている、心理学的見地に基づき感情を 22 種類に分類した理論である、OCC theory[4] について述べる。

OCC theory は、Ortony, Clore, Collins らが提唱した、感情をモデル化した理論であり、人間の包括的な感情をモデル化しており、感情の特徴付けが明確であるために、比較的理解しやすい。そのため、計算機科学の分野において広く用いられている。

OCC theory によって分類された 22 種類の感情を以下に示す。

1. 事象の結果の望ましさに関して
    - (a) 自分にとっての結果の望ましさに関して
        - i. イベントに関して (Well-being Emotions)
          Joy(喜び), Distress(嘆き)
        - ii. 予想に関して (Prospected-based Emotions)
            - A. 単なる予想に関して (Prospect Emotions)
              Hope(望み), Fear(恐れ)
            - B. 予想していたことが実際に起こったかどうかに関して (Confirmation Emotions)
                - ① 予想していたことが実際に起こったことに関して
                  Satisfaction(満足), FearConfirmed(恐れていたことが確定)
                - ② 予想していたことが実際に起こらなかったことに関して
                  Relief(安堵), Disappointment(落胆)
    - (b) 他者にとっての結果の望ましさに関して (Fortunes-of-others Emotions)
        - i. 他者が良い結果を得る
          HappyFor(共に喜ぶ), Resentment(憤り)
        - ii. 他者が悪い結果を得る
          SorryFor(共に残念に思う), Gloating(ほくそ笑む)
2. 行動に対する賞賛度に対して
    - (a) 行動に関して (Attribution Emotions)
        - i. 自分の行動に関して
          Pride(自尊心), Shame(羞恥心)

ii. 他者の行動に関して
Admiration(称賛), Reproach(非難)

(b) 行動と起こった事象に関して (Well-being/Attribution Emotions)

i. 自分の行動と起こった事象に関して
Gratification(達成感), Remorse(後悔)

ii. 他者の行動と起こった事象に関して
Gratitude(謝意), Anger(怒り)

3. 対象に対する好き嫌い (Attraction Emotions)
Love(好き), Hate(嫌い)

3 先行研究

3.1 Adam らの研究

Adam ら [1] は、エージェントの感情表現を可能にするため、OCC theory と BDi モデルを用いている。そして、BDI モデルでの形式化に用いられる論理体系 BDI logic に対し、新たなオペレータを導入し拡張することで、OCC theory の Attraction Emotions の感情を除く 20 種類の感情を形式化している。ここで、Attraction Emotions の感情の形式化を省略しているのは、Adam らの論理体系では、対象物を引数としてとるために必要な一階述語論理を扱えないためである。

また、Adam らの形式化では、感情の生起条件にのみ焦点を当てており、OCC theory で定義されている、感情の強さに影響する変数については考慮されていない。そのため、感情は生起されているかどうかしか扱えず、感情の強さについては扱えない。それに加えて、信念の変更について考慮されていないため、時間経過による信念の変化は扱えない。

しかし、後述の Stunebrink ら [2] の研究に比べ、用いるオペレータが少なく、エージェントの感情表現を実装するには適している。そのため Adam らの形式化を基に、より OCC theory に忠実になるように拡張を行った。

3.2 Steunebrink らの研究

Steunebrink[2] らは、Adam らの形式化よりも、より OCC theory に忠実に形式化を行っている。OCC theory の感情の生起条件と、感情の強さに影響を与える変数を基に、Adam らよりも多くのオペレータを用いている。特に、感情を引き起こすトリガーについて注目しており、そのトリガーとして、「イベントの結果の認識」、「エージェントのアクションの認識」、「オブジェクトの認識」等をオペレータとして、導入している。また、それらのオペレータを用いて、「何かを認識する」というオペレータも導入されている。これらのオペレータを用いて、例えば、認識したものが「イベントの結果」であるなら、その結果は予想される結果であるか、実際の結果であるかの判別によって、生起される感情が区別される。これは、OCC theory でも述べられている、感情タイプの分類と一致しており、OCC theory に忠実な形式化であると言える。

それに加え、信念の変更を可能とするオペレータも導入されている。そのため、感情が生起しているかどうかではなく、感情が「今」生起したかどうかを表現することが可能であり、人間らしい感情表現であると言える。また、これは、本論文で我々が提案する時間経過による感情の強さの減衰を可能とする論理体系を構築する上で、非常に重要視するべきことである。

しかし、実装をするにあたって、用いるオペレータが非常に多くなってしまうのは難点であるため、Adam らの形式化を基に、Steunebrink らの形式化に用いられた信念の変更を取り入れることにした。

4　構文

本章では、我々が提案する論理体系での論理式の定義を与える。我々の論理体系は、LTL(線形時相論理) に感情表現の定義に用いるさまざまな様相オペレータを導入し、さらに述語論理に拡張したものである。

まず、以下のものを定めておく。

- 命題記号の無限集合を $ATM = \{P_1, P_2, \ldots\}$ とする。また、命題記号は論理式である。
- エージェントを表す記号の無限集合を $AGT = \{i_1, i_2, \ldots\}$ とする。
- 述語記号の無限集合を $PRD = \{R_1, R_2, \ldots\}$ とする。
  また、$Appealing, Unaapealing, Familiar \in PRD$ とする。
- 定数記号の無限集合を $CST = \{c_1, c_2, \ldots\}$ とする。
- 変数記号の無限集合を $VRB = \{x_1, x_2, \ldots\}$ とする。
- 関数記号の無限集合を $FUC = \{f_1, f_2, \ldots\}$ とする。

以上を定めておいたうえで、項および論理式を以下のように定義する。

- $c \in CST$ ならば、c は項である。
- $x \in VRB$ ならば、x は項である。
- $f \in FUC$、$t_1, t_2, ..., t_n (n > 0)$ が項ならば、$f(t_1, t_2, ..., t_n)$ は項である。
- $r \in PRD$、$t_1, t_2, ..., t_n (n \geq 0)$ が項ならば、$r(t_1, t_2, ..., t_n)$ は論理式である。これを原始論理式と呼ぶ。
- $x \in VRB$、$\varphi(x)$ が論理式ならば、$\forall x \varphi(x)$、$\exists x \psi(x)$ は論理式である。
- φ_1, φ_2 が論理式ならば、$\neg\varphi_1, \varphi_1 \vee \varphi_2$ も論理式であるとする。また、$\wedge, \rightarrow, \leftrightarrow$ の省略形を以下のように定義する。
  - $\varphi_1 \wedge \varphi_2 := \neg(\neg\varphi_1 \vee \neg\varphi_2)$
  - $\varphi_1 \rightarrow \varphi_2 := \neg\varphi_1 \vee \varphi_2$

- $\varphi_1 \leftrightarrow \varphi_2 := (\varphi_1 \rightarrow \varphi_2) \wedge (\varphi_2 \rightarrow \varphi_1)$

- φ_1, φ_2 が論理式ならば、$\neg\varphi_1, \varphi_1 \vee \varphi_2, \varphi_1 U \varphi_2$ も論理式である。

- φ が論理式、$i \in AGT$、d(desirability) が $0 \leq d \leq 1$ を満たす実数の時、$Des^i_d \varphi$ は論理式である。
  直感的には、$Des^i_d \varphi$ は「エージェント i にとって φ は度合い d で望ましい」ことを表す。

- φ が論理式、$i \in AGT$、l(likelihood) が $0 \leq l \leq 1$ を満たす実数の時、$Prob^i_l \varphi$ は論理式である。
  省略形として、以下を定義する。

$$Bel^i \varphi := Prob^i_{1.0} \varphi$$

  直感的には、$Prob^i_l \varphi$ は「エージェント i は φ が見込み l で真になる可能性があると信じている」、$Bel^i \varphi$ は「エージェント i は φ が真であると信じている」ことを表す。

- φ が論理式ならば、$G\varphi$ は論理式である。
  省略形として、以下を定義する。

$$F\varphi := \neg G \neg \varphi$$

  直感的には、$G\varphi$ は「φ は現在時刻以降ずっと成り立つ」、$F\varphi$ は「φ は現在時刻以降のある時点で成り立つ」ことを表す。

- φ が論理式ならば、$H\varphi$ は論理式である。
  省略形として、以下を定義する。

$$P\varphi := \neg H \neg \varphi$$

  直感的には、$H\varphi$ は「φ は現在時刻までずっと成り立つ」、$P\varphi$ は「φ は現在時刻までのある時点で成り立つ」ことを表す。

- φ が論理式ならば、$X\varphi$ は論理式である。
  直感的には、$X\varphi$ は「φ は現在時刻より 1 時刻後で成り立つ」ことを表す。

- φ が論理式ならば、$Y\varphi$ は論理式である。
  省略形として、以下を定義する。

  - $Bel\_arise^i \varphi := Bel^i \varphi \wedge \neg Y Bel^i \varphi \quad (1)$

  - $Prob\_arise^i_l \varphi := Prob^i_l \varphi \wedge Y Prob^i_0 \varphi \quad (2)$

  直感的には、$Y\varphi$ は「φ は現在時刻より 1 時刻前で成り立つ」ことを表す。

- φ が論理式、$i \in AGT$、e が $0 \leq e \leq 1$ を満たす実数の時、$Effort^i_e \varphi$ は論理式である。
  直感的には、$Effort^i_e \varphi$ は「エージェント i は φ を真にするため、e 程度努力した」ことを表す。

- φ が論理式、$i \in AGT$、v(value) が $0 \leq v \leq 1$ を満たす実数の時、$Deserve^i_v\varphi$ は論理式である。
  直感的には、$Deserve^i_v\varphi$ は「エージェント i にとって φ が程度 v で相応しい」ことを表す。

- φ が論理式、$i \in AGT$、p(praiseworthiness) が $0 \leq p \leq 1$ を満たす実数の時、$Praise^i_p\varphi$ は論理式である。
  直感的には、$Praise^i_d\varphi$ は「エージェント i が φ を真にすることは、程度 p で評価される」ことを表す。

この他、$\wedge, \vee, \neg$ などのオペレータ間に一般的な優先順位を導入する。
　述語記号、定数記号、変数記号、関数記号を組にしたものを一階言語と呼ぶことにする。一階言語を定めると、一階言語の原始論理式の集合が決まるのでこれを AF と書く。
　ただし、述語記号 $Apealing, Unappealing, Familiar$ については、以下のように表す。

- $Appealing$ については、任意の $i \in AGT, j \in AGT$ と、$0 \leq a \leq 1$ を満たす実数 a に対し、$Apealing(i, j, a)$ を原始論理式とする。直感的には、$Appealing(i, j, a)$ は、エージェント i がエージェント j に対し、程度 a の好意を感じることを表す。

- $Unappealing$ については、任意の $i \in AGT, j \in AGT$ と、$0 \leq a \leq 1$ を満たす実数 a に対し、$Unapealing(i, j, a)$ を原始論理式とする。直感的には、$Unappealing(i, j, a)$ は、エージェント i がエージェント j に対し、程度 a の嫌悪を感じることを表す。

- $Familiar$ については、任意の $i \in AGT, j \in AGT$ と、$0 \leq f \leq 1$ を満たす実数 f に対し、$Familiar(i, j, f)$ を原始論理式とする。直感的には、$Familiar(i, j, f)$ は、エージェント i がエージェント j に対し、程度 f の親しさを感じることを表す。

5 意味論

5.1 構造

　領域の無限集合を $U = \{u_1, u_2, \ldots\}$ とし、以後、U の要素は全て CST に属するものとする。また、以下、$\mathbb{B} = \{\top, \bot\}$ とする。
　以下のものを定めておく。

- 可能世界の集合 $W (\neq \emptyset)$

- 各世界 w ごとに、領域 U のもとでの一階述語論理の解釈を１つ決めると、w での原始論理式の真偽が決まる。
  これを関数
  $$V : W \times AF \to \mathbb{B}$$
  で表す。ただし、$Appealing, Unappealing, Familiar$ については、以下のように決めておく。

- *Appealing* については、各世界 w ごとに、任意の i, j に対し、$0 \leq a \leq 1$ を満たし、$Appealing(i, j, a)$ を真にする a が一つだけ存在することとする。
- *Unappealing* については、各世界 w ごとに、任意の i, j に対し、$0 \leq a \leq 1$ を満たす $Appealing(i, j, a)$ が存在する時、$Unappealing(i, j, 1-a)$ が存在することとする。
- *Familiar* については、各世界ごとに、w 任意の i, j に対し、$0 \leq f \leq 1$ を満たし、$Familiar(i, j, f)$ を真にする f が一つだけ存在することとする。

- AGT の各要素 i に対し、$W \times W$ から [0,1] への関数

$$\mathscr{B}^i : W \times W \to [0,1]$$

ただし任意の $i \in AGT$、$w \in W$ に対し、$\sum_{w' \in W} \mathscr{B}^i(w, w') = 1$ であること。また、エージェントが自身の心的状態に関して完全な信念を持つという、通常の BDI logic と同じ性質を成り立たせるため、$\mathscr{B}^i$ は以下を満たすものとする。

$$\mathscr{B}^i(w, w') > 0 \text{ ならば、} \mathscr{B}^i(w, w'') = \mathscr{B}^i(w', w'')$$

$\mathscr{B}^i$ は、エージェント i にとっての度合い付きの信念を表す関数である。
以下、$\mathscr{B}^i(w, w') > 0$ が成り立つとき、w' は w から $\mathscr{B}^i$ で到達可能な世界であると言う。
また、ある世界 w において、$Appealing, Unappealing, Familiar$ が成り立つのであれば、w から $\mathscr{B}^i$ で到達可能な世界 w' においても同様に成り立つこととする。

- AGT の各要素 i に対し、W から [0,1] への関数

$$\mathscr{D}^i : W \to [0,1]$$

エージェント i にとっての各世界の望ましさを定める関数である。

- 可能世界の集合 W から W への全単射

$$\mathscr{X} : W \to W$$

この $\mathscr{X}$ は可能世界 w の未来の可能世界 $\mathscr{X}(w)$ を表す関数である。同様に、$\mathscr{X}^{-1}$ も可能世界 w の過去の可能世界 $\mathscr{X}^{-1}(w)$ を表す関数である。
ただし、任意の 0 以上の整数 t に対して、$\mathscr{B}^i(w, w_1) > 0$ かつ、$w_2 = \mathscr{X}^{-t}(w)$ かつ、$w_3 = \mathscr{X}^{-t}(w_1)$ かつ、$\mathscr{B}^i(w_2, w') = d > 0$ ならば、$\mathscr{B}^i(w_3, w') = d$ を満たすものとする。この制約を (1P) と呼ぶことにする。
ここで、$\mathscr{X}^t(w)$ は、w に関数 $\mathscr{X}$ を t 回適用して得られる世界で、w の t 時刻後の世界、$\mathscr{X}^{-t}(w)$ は、w に関数 $\mathscr{X}^{-1}$ を t 回適用して得られる世界で、w の t 時刻前の世界である。

- AGT の各要素 i に対し、$w' \in \mathscr{G}(w)$ を満たす w と w' の組から、[0,1] への関数

$$\mathscr{E}^i : \{(w, w') | w' \in \mathscr{G}(w)\} \to [0,1]$$

エージェント i の努力を表す関数である。ここで、関数 $\mathscr{G}, \mathscr{H}$ を W から 2^W への関数とし、以下のように定める。

- $\mathscr{G}(w) = \{\mathscr{X}^n(w) \mid n$ は非負整数 $\}$

- $\mathscr{H}(w) = \{\mathscr{X}^{-n}(w) \mid n$ は非負整数 $\}$

- AGT の各要素 i に対し、W から [0,1] への関数

$$\mathscr{V}^i : W \to [0,1]$$

エージェント i にとっての各世界の相応しさを定める関数である。

- AGT の各要素 i に対し、W から [0,1] への関数

$$\mathscr{P}^i : W \to [0,1]$$

エージェント i にとっての各世界の称賛度を定める関数である。

以上を組にしたものを構造 M と呼ぶ。

図 1 は、先に述べた制約 (1P) を表した図である。これは、現在時刻 w の 1 時刻前の世界 $w_3(\mathscr{X}^{-1}(w))$ と、現在時刻 w から $\mathscr{B}^i$ で到達可能な世界 $w_1(\mathscr{B}^i(w))$ の 1 時刻前の世界 $w_2(\mathscr{X}^{-1}(w))$ から、$\mathscr{B}^i$ で到達可能な世界は同じ世界であるということである。

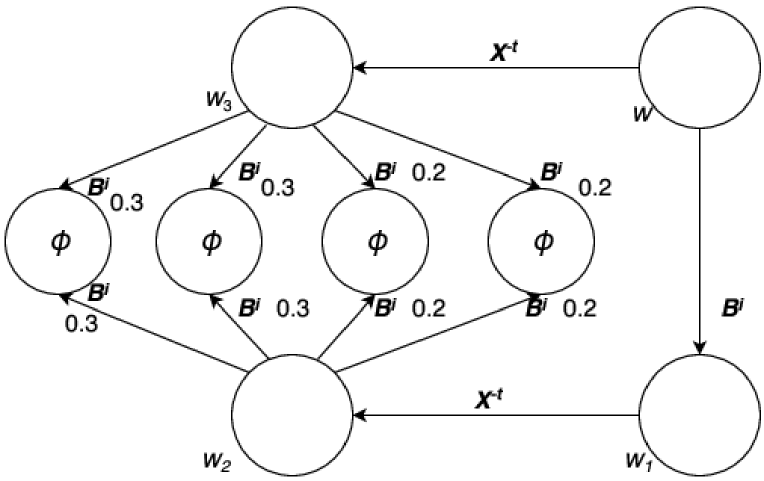

図 1: 制約 (1P)

5.2 解釈

論理式 φ と構造 M、および M の世界 w に対し、φ の M, w での解釈を $[\![\varphi]\!]_{\langle M,w\rangle}$ を以下のように定義する。

- $P \in AF$ ならば、$[\![P]\!]_{\langle M,w\rangle} = \top$ **iff** $V(w,P) = \top$

- $[\![\neg\varphi]\!]_{\langle M,w\rangle} = \top$ **iff not** $[\![\varphi]\!]_{\langle M,w\rangle} = \top$

- $[\![\varphi \vee \psi]\!]_{\langle M,w\rangle} = \top$ **iff** $[\![\varphi]\!]_{\langle M,w\rangle} = \top$ **or** $[\![\psi]\!]_{\langle M,w\rangle} = \top$
- $[\![\forall x\ \varphi(x)]\!]_{\langle M,w\rangle} = \top$ **iff** $u \in U$ を満たす全ての u に対し、 $[\![\varphi(u)]\!]_{\langle M,w\rangle} = \top$
- $[\![\exists x\ \varphi(x)]\!]_{\langle M,w\rangle} = \top$ **iff** $u \in U$ を満たす、ある u に対し、 $[\![\varphi(u)]\!]_{\langle M,w\rangle} = \top$
- $[\![Prob^i_l \varphi]\!]_{\langle M,w\rangle} = \top$ **iff** $\sum_{\substack{w' \in W \text{ かつ} \\ w' \text{で} \varphi \text{ が真}}} \mathscr{B}^i(w,w') = l$
- $[\![Effort^i_e \varphi]\!]_{\langle M,w\rangle} = \top$ **iff** $w \in \mathscr{G}(w')$ を満たす w' が存在して、
  $[\![\varphi]\!]_{\langle M,w'\rangle} = \top$ **and** $\mathscr{E}^i(w,w') = e$
- $[\![Des^i_d \varphi]\!]_{\langle M,w\rangle} = \top$ **iff** $\dfrac{\sum_{\substack{w' \in W \text{ かつ} \\ w' \text{で} \varphi \text{ が真}}} (\mathscr{D}^i(w') \times \mathscr{B}^i(w,w'))}{\sum_{\substack{w' \in W \text{ かつ} \\ w' \text{で} \varphi \text{ が真}}} \mathscr{B}^i(w,w')} = d$
- $[\![Deserve^i_v \varphi]\!]_{\langle M,w\rangle} = \top$ **iff** $\dfrac{\sum_{\substack{w' \in W \text{ かつ} \\ w' \text{で} \varphi \text{ が真}}} (\mathscr{V}^i(w') \times \mathscr{B}^i(w,w'))}{\sum_{\substack{w' \in W \text{ かつ} \\ w' \text{で} \varphi \text{ が真}}} \mathscr{B}^i(w,w')} = v$
- $[\![Praise^i_p \varphi]\!]_{\langle M,w\rangle} = \top$ **iff** $\dfrac{\sum_{\substack{w' \in W \text{ かつ} \\ w' \text{で} \varphi \text{ が真}}} (\mathscr{P}^i(w') \times \mathscr{B}^i(w,w'))}{\sum_{\substack{w' \in W \text{ かつ} \\ w' \text{で} \varphi \text{ が真}}} \mathscr{B}^i(w,w')} = p$
- $[\![G\phi]\!]_{\langle M,w\rangle} = \top$ **iff** 任意の $w' \in \mathscr{G}(w)$ に対して $[\![\phi]\!]_{\langle M,w'\rangle} = \top$
- $[\![H\phi]\!]_{\langle M,w\rangle} = \top$ **iff** 任意の $w' \in \mathscr{H}(w)$ に対して $[\![\phi]\!]_{\langle M,w'\rangle} = \top$
- $[\![X\phi]\!]_{\langle M,w\rangle} = \top$ **iff** $[\![\phi]\!]_{\langle M,\mathscr{X}(w)\rangle} = \top$
- $[\![Y\phi]\!]_{\langle M,w\rangle} = \top$ **iff** $[\![\phi]\!]_{\langle M,\mathscr{X}^{-1}(w)\rangle} = \top$

任意の構造 M の任意の世界 w に対し、$[\![\varphi]\!]_{\langle M,w\rangle} = \top$ が成り立つ時、φ が恒真であるといい、$\vdash \varphi$ と書く。

6 恒真論理式

我々の論理体系では、いくつかの基本的な性質を満たす。ここで示す性質は、付録に記述する証明の中で使用する。以下で、単に論理式が掲げてあるものは、その論理式が恒真であるという性質を示す。

$\Box$ を Bel^i, G, H のうち、いずれか１つのオペレータとするとき、以下が成り立つ。

$$\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi) \tag{K-$\Box$}$$

$$\frac{\varphi}{\Box\varphi} \tag{RN-$\Box$}$$

$$\Box(\varphi \wedge \psi) \leftrightarrow (\Box\varphi \wedge \Box\psi) \tag{C-$\Box$}$$

ただし、$\frac{\varphi}{\psi}$ は、「φ が恒真ならば、ψ も恒真である」という性質を表す。

また、$\Diamond$ を、F, P のうち、いずれか１つのオペレータとするとき、以下が成り立つ。

$$\frac{\varphi}{\Diamond\varphi} \tag{RN-$\Diamond$}$$

$$\Diamond(\varphi \wedge \psi) \rightarrow (\Diamond\varphi \wedge \Diamond\psi) \tag{C-$\Diamond$}$$

オペレータ Bel^i は公理系 KD45 に当てはまる。このオペレータは、一般的な様相論理の公理に加え、以下の論理式を恒真にする。

$$Bel^i\varphi \rightarrow \neg Bel^i\neg\varphi \tag{D-Bel^i}$$

$$Bel^i\varphi \rightarrow Bel^i Bel^i\varphi \tag{4-Bel^i}$$

$$\neg Bel^i\varphi \rightarrow Bel^i\neg Bel^i\varphi \tag{5-Bel^i}$$

また、我々の論理体系で新たに導入したオペレータ Bel_arise^i においても、上記と同様の論理式を恒真にする。これらの証明については、付録で述べる。

$$Bel\_arise^i\varphi \rightarrow \neg Bel\_arise^i\neg\varphi \tag{D-Bel_arise}$$

$$Bel\_arise^i\varphi \rightarrow Bel^i Bel\_arise^i\varphi \tag{4-Bel_arise^i}$$

$$\neg Bel\_arise^i\varphi \rightarrow Bel^i\neg Bel\_arise^i\varphi \tag{5-Bel_arise^i}$$

以上の恒真論理式は、信念に矛盾がない (D-Bel^i)(D-Bel_arise^i)、自身の信念を自覚している (4-Bel^i)(4-Bel_arise^i)、自身の信念でないものを信念でないものと自覚している (5-Bel_arise^i) ことを示している。

$\Box$ が、オペレータ $Prob^i_l, Des^i_d, Praise^i_p, Deserve^i_v$ のいずれかの場合、以下が論理式が成り立つ。自身の $Prob^i_l$ などを自覚していることを示している。

$$\Box\varphi \leftrightarrow Bel^i\Box\varphi \tag{4-MIX}$$

$$\neg\Box\varphi \leftrightarrow Bel^i\neg\Box\varphi \tag{5-MIX}$$

7　感情

本章では、OCC theory で定義されている感情を、Adam らの形式化を踏まえて、4 章で定義したオペレータを用いて定義する。我々の従来研究では、ある感情の生起を状況に関する、ある信念の存在を基に定義していた。しかし、これではその信念が存続する限り、同様に感情が生起することになり、感情の時間経過による減衰を表現することは困難である。そこで、本論文では 4 章で略記として導入した、信念が生起したことを表すオペレータ Bel_arise を用いて、ある感情が生起したことを表す論理式を定義し、それを用いて、感情が時間経過により減衰することを表す論理式を定義する。

また、OCC theory で定義されている 22 個の感情は、以下の 7 グループに分類され、同じグループに属する感情は、生起条件が同様の形で定義されている。以下の各分類の先頭の記号は、2 節の分類に対応している。

1(a)i　Well-being emotions：Joy，Distress

1b Fortunes-of-others emotions：
Happy-For，Sorry-For， Resentment，Gloating

1(a)iiA Prospect emotions：Hope，Fear

1(a)iiB Confirmation emotions：
Satisfaction，Fear-Confirmed，Relief，Disappointment

2a Attribution emotions：
(Self-focued)：Pride，Shame
(Other-focued)：Admiration，Reproach

2b Well-being/Attribution compounds emotions：
(Self-focued)：Gratification，Remorse
(Other-focued)：Gratitude，Anger

2 Attraction emotions：Love，Hate

7.1 Well-being emotions

Joy

Joy の生起条件は、従来研究 [9] では OCC theory に沿って「ある事柄 φ の生起を信じ、かつそれが望ましい事柄であるとき、その事柄に関する Joy の生起が生起する」としていた。我々もこれを踏襲するが、我々は、感情の時間経過による減衰を考慮するため、以下のような定義を行う。他のグループの感情についても同様である。

$$Joy\_arise^{i}_{f_{J'}(d)}\ \varphi := Des^{i}_{d}\ \varphi\ \wedge\ Bel\_arise^{i}\ \varphi$$

$Joy^{i}_{f_J(d,t)}\varphi$ を新たなオペレータとし、$Joy^{i}_{f_J(d,t)}\varphi$ の解釈を以下のように定義する。

$$[\![Joy^{i}_{f_J(d,t)}\varphi]\!]_{\langle M,w\rangle} = \top \textbf{ iff } [\![Joy\_arise^{i}_{d}\varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top$$

ただし、t は 0 以上の整数とし、上記の iff の右側を満たすような t が複数存在する場合、最小の t をとることにする。また、関数 $f_{J'}$ は、d に関する増加関数であり、関数 f_J は、t に関する減少関数で、$f_J(d,0) = d$ を満たすものとする。関数 f_J、関数 $f_{J'}$ のどちらも、値域は $[0,1]$ とする。

Distress

Distress の生起条件を以下のように定義する。

$$Distress\_arise^{i}_{f_{D'}(d)}\ \varphi := Des^{i}_{d}\ \varphi\ \wedge\ Bel\_arise^{i}\ \varphi$$

$Distress^{i}_{f_D(d,t)}\varphi$ を新たなオペレータとし、$Distress^{i}_{f_D(d,t)}\varphi$ の解釈を以下のように定義する。

$$[\![Distress^{i}_{f_D(d,t)}\varphi]\!]_{\langle M,w\rangle} = \top \textbf{ iff } [\![Distress\_arise^{i}_{d}\varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top$$

ただし、t は 0 以上の整数とし、上記の iff の右側を満たすような t が複数存在する場合、最小の t をとることにする。また、関数 $f_{D'}$ は、d に関する減少関数であり、関数 f_D は、t に関する減少関数で、$f_D(d,0) = d$ を満たすものとする。関数 f_D、関数 $f_{D'}$ のどちらも、値域は $[0,1]$ とする。

7.2 Fortune-of-others emotions

ここで出てくる $Love, Hate$ の定義は、後の 7.7 章で述べている。

HappyFor

OCC theory において、HappyFor は、「他者にとって望ましいと推測されるイベントについて喜ぶ」感情タイプと定義されており、その感情の強さに影響する変数には、以下の変数が定義されている。

- 他者にとって望ましいイベントがどの程度自分にとって望ましいか
- 他者にとってイベントがどの程度望ましいと推測されるか
- 他者がどの程度イベントに相応しいか
- 自分が他者をどの程度好ましく思っているか

よって、HappyFor の生起条件を以下のように定義する。

$$\begin{aligned} \mathit{HappyFor\_arise}^{i,j}_{f_{HF'}(d_i,d_j,v,LD)}\ \varphi := \mathit{Bel\_arise}^i\ \varphi\ \wedge\ \mathit{Des}^i_{d_i}\mathit{Bel}^j\ \varphi\ \wedge\ \mathit{Bel}^i\mathit{Des}^j_{d_j}\ \varphi \\ \wedge\ \mathit{Bel}^i\mathit{Deserve}^j_v\ \varphi\ \wedge\ \mathit{Love}^i_{LD}\, j \end{aligned}$$

$\mathit{HappyFor}^{i,j}_{f_{HF}(d,t)}\varphi$ を新たなオペレータとし、$\mathit{HappyFor}^{i,j}_{f_{HF}(d,t)}\varphi$ の解釈を以下のように定義する。

$$[\![\mathit{HappyFor}^{i,j}_{f_{HF}(d,t)}\varphi]\!]_{\langle M,w\rangle} = \top \textbf{ iff } [\![\mathit{HappyFor\_arise}^{i,j}_{d}\varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top$$

ただし、t は 0 以上の整数とし、上記の iff の右側を満たすような t が複数存在する場合、最小の t をとることにする。また、関数 $f_{HF'}$ は、d_i, d_j, v, LD に関する増加関数、関数 f_{HF} は、t に関する減少関数で、$f_{HF}(d,0) = d$ を満たすものとする。関数 f_{HF}、関数 $f_{HF'}$ のどちらも、値域は $[0,1]$ とする。

SorryFor

OCC theory において、SorryFor は、「他者にとって望ましくないと推測されるイベントについて不快に思う」感情タイプと定義されており、その感情の強さに影響する変数には、以下の変数が定義されている。

- 他者にとって望ましくないイベントがどの程度自分にとって望ましくないか

- 他者にとってイベントがどの程度望ましくないと推測されるか
- 他者がどの程度イベントに相応しくないか
- 自分が他者をどの程度好ましく思っているか

よって、SorryFor の生起条件を以下のように定義する。

$$SorryFor\_arise^{i,j}_{f_{HF'}(d_i,d_j,v,LD)}\ \varphi := Bel\_arise^{i}\ \varphi\ \wedge\ Des^{i}_{d_i}Bel^{j}\ \varphi\ \wedge\ Bel^{i}Des^{j}_{d_j}\ \varphi \\ \wedge\ Bel^{i}Deserve^{j}_{v}\ \varphi\ \wedge\ Love^{i}_{LD}j$$

$SorryFor^{i,j}_{f_{SF}(d,t)}\varphi$ を新たなオペレータとし、$SorryFor^{i,j}_{f_{SF}(d,t)}\varphi$ の解釈を以下のように定義する。

$$[\![SorryFor^{i,j}_{f_{HF}(d,t)}\varphi]\!]_{\langle M,w\rangle} = \top \textbf{ iff } [\![SorryFor\_arise^{i,j}_{d}\varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top$$

ただし、t は 0 以上の整数とし、上記の iff の右側を満たすような t が複数存在する場合、最小の t をとることにする。また、関数 $f_{SF'}$ は、LD に関する増加関数、d_i, d_j, v に関する減少関数、関数 f_{SF} は、t に関する減少関数で、$f_{SF}(d,0) = d$ を満たすものとする。関数 f_{SF}、関数 $f_{SF'}$ のどちらも、値域は $[0,1]$ とする。

Resentment

OCC theory において、Resentment は、「他者にとって望ましいと推測されるイベントについて不快に思う」感情タイプと定義されており、その感情の強さに影響する変数には、以下の変数が定義されている。

- 他者にとって望ましいイベントがどの程度自分にとって望ましくないか
- 他者にとってイベントがどの程度望ましいと推測されるか
- 他者がどの程度イベントに相応しくないか
- 自分が他者をどの程度好ましく思ってないか

よって、Resentment の生起条件を以下のように定義する。

$$Resentment\_arise^{i,j}_{f_{Rs'}(d_i,d_j,v,HTD)}\ \varphi := Bel\_arise^{i}\ \varphi\ \wedge\ Des^{i}_{d_i}Bel^{j}\ \varphi\ \wedge\ Bel^{i}Des^{j}_{d_j}\ \varphi \\ \wedge\ Bel^{i}Deserve^{j}_{v}\ \varphi\ \wedge\ Hate^{i}_{HTD}j$$

$Resentment^{i,j}_{f_{Rs}(d,t)}\varphi$ を新たなオペレータとし、$Resentment^{i,j}_{f_{Rs}(d,t)}\varphi$ の解釈を以下のように定義する。

$$[\![Resentment^{i,j}_{f_{Rs}(d,t)}\varphi]\!]_{\langle M,w\rangle} = \top\textbf{iff}\quad [\![Resentment\_arise^{i,j}_{d}\varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top$$

ただし、t は 0 以上の整数とし、上記の iff の右側を満たすような t が複数存在する場合、最小の t をとることにする。また、関数 $f_{Rs'}$ は、d_j, HTD に関する増加関数、d_i, v に関する減少関数、関数 f_{Rs} は、t に関する減少関数で、$f_{Rs}(d,0) = d$ を満たすものとする。関数 f_{Rs}、関数 $f_{Rs'}$ のどちらも、値域は $[0,1]$ とする。

Gloating

OCC theory において、Gloating は、「他者にとって望ましくないと推測されるイベントについて喜ぶ」感情タイプと定義されており、その感情の強さに影響する変数には、以下の変数が定義されている。

- 他者にとって望ましくないイベントがどの程度自分にとって望ましいか
- 他者にとってイベントがどの程度望ましくないと推測されるか
- 他者がどの程度イベントに相応しいか
- 自分が他者をどの程度好ましく思ってないか

よって、Gloating の生起条件を以下のように定義する。

$$\begin{aligned} \mathit{Gloating\_arise}^{i,j}_{f_{Gl'}(d_i,d_j,v,HTD)}\ \varphi := \mathit{Bel\_arise}^i\ \varphi\ \wedge\ \mathit{Des}^i_{d_i}\mathit{Bel}^j\ \varphi\ \wedge\ \mathit{Bel}^i\mathit{Des}^j_{d_j}\ \varphi \\ \wedge\ \mathit{Bel}^i\mathit{Deserve}^j_v\ \varphi\ \wedge\ \mathit{Hate}^i_{HTD}j \end{aligned}$$

$\mathit{Gloating}^{i,j}_{f_{Gl}(d,t)}\varphi$ を新たなオペレータとし、$\mathit{Gloating}^{i,j}_{f_{Gl}(d,t)}\varphi$ の解釈を以下のように定義する。

$$[\![\mathit{Gloating}^{i,j}_{f_{Gl}(d,t)}\varphi]\!]_{\langle M,w\rangle} = \top \textbf{iff} \quad [\![\mathit{Gloating\_arise}^{i,j}_{d}\varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top$$

ただし、t は 0 以上の整数とし、上記の iff の右側を満たすような t が複数存在する場合、最小の t をとることにする。また、関数 $f_{Gl'}$ は、d_i, v, HTD に関する増加関数、d_j に関する減少関数、関数 f_{Gl} は、t に関する減少関数で、$f_{Gl}(d,0) = d$ を満たすものとする。関数 f_{Gl}、関数 $f_{Gl'}$ のどちらも、値域は $[0,1]$ とする。

7.3 Prospect emotions

Hope

Hope の生起条件を以下のように定義する。

$$\mathit{Hope\_arise}^i_{f_{H'}(d,l)}\ \varphi := \mathit{Des}^i_d\ \varphi\ \wedge\ \mathit{Prob\_arise}^i_l\ \varphi\ \wedge\ l \neq 1$$

$\mathit{Hope}^i_{f_H(d,t)}\varphi$ を新たなオペレータとし、$\mathit{Hope}^i_{f_H(d,t)}\varphi$ の解釈を以下のように定義する。

$$[\![\mathit{Hope}^i_{f_H(d,t)}\varphi]\!]_{\langle M,w\rangle} = \top\ \textbf{iff}\ [\![\mathit{Hope\_arise}^i_d\varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top$$

ただし、t は 0 以上の整数とし、上記の iff の右側を満たすような t が複数存在する場合、最小の t をとることにする。また、関数 $f_{H'}$ は、d, l に関する増加関数、関数 f_H は、t に関する減少関数で、$f_H(d,0) = d$ を満たすものとする。関数 f_H、関数 $f_{H'}$ のどちらも、値域は $[0,1]$ とする。

Fear

Fear の生起条件を以下のように定義する。

$$Fear\_arise^{i}_{f_{F'}(d,l)}\ \varphi := Des^{i}_{d}\ \varphi \wedge Prob\_arise^{i}_{l}\ \varphi \wedge l \neq 1$$

$Fear^{i}_{f_F(d,t)}\varphi$ を新たなオペレータとし、$Fear^{i}_{f_F(d,t)}\varphi$ の解釈を以下のように定義する。

$$[\![Fear^{i}_{f_F(d,t)}\varphi]\!]_{\langle M,w\rangle} = \top \text{ } \mathbf{iff} \text{ } [\![Fear\_arise^{i}_{d}\varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top$$

ただし、t は 0 以上の整数とし、上記の iff の右側を満たすような t が複数存在する場合、最小の t をとることにする。また、関数 $f_{F'}$ は、d に関する減少関数、l に関する増加関数、関数 f_F は、t に関する減少関数で、$f_F(d,0) = d$ を満たすものとする。関数 f_F、関数 $f_{F'}$ のどちらも、値域は $[0,1]$ とする。

7.4 Comfirmation emotions

Satisfaction

Satisfaction の生起条件を以下のように定義する。

$$Satisfaction\_arise^{i}_{f_{St'}(x,e,l)}\ \varphi := Bel^{i}PHope^{i}_{x}\ \varphi \wedge Bel^{i}PEffort^{i}_{e}\varphi \wedge Prob\_arise^{i}_{l}\ \varphi$$

$Satisfaction^{i}_{f_{St}(d,t)}\varphi$ を新たなオペレータとし、$Satisfaction^{i}_{f_{St}(d,t)}\varphi$ の解釈を以下のように定義する。

$$[\![Satisfaction^{i}_{f_{St}(d,t)}\varphi]\!]_{\langle M,w\rangle} = \top \text{ } \mathbf{iff} \text{ } [\![Satisfaction\_arise^{i}_{d}\varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top$$

ただし、t は 0 以上の整数とし、上記の iff の右側を満たすような t が複数存在する場合、最小の t をとることにする。また、関数 $f_{St'}$ は、x,e,l に関する増加関数、関数 f_{St} は、t に関する減少関数で、$f_{St}(d,0) = d$ を満たすものとする。関数 f_{St}、関数 $f_{St'}$ のどちらも、値域は $[0,1]$ とする。

FearConfirmed

FearConfirmed の生起条件を以下のように定義する。

$$FearConfirmed\_arise^{i}_{f_{FC'}(x,e,l)}\ \varphi := Bel^{i}PFear^{i}_{x}\ \varphi \wedge Bel^{i}PEffort^{i}_{e}\neg\varphi \wedge Prob\_arise^{i}_{l}\ \varphi$$

$FearConfirmed^{i}_{f_{FC}(d,t)}\varphi$ を新たなオペレータとし、$FearConfirmed^{i}_{f_{FC}(d,t)}\varphi$ の解釈を以下のように定義する。

$$[\![FearConfirmed^{i}_{f_{FC}(d,t)}\varphi]\!]_{\langle M,w\rangle} = \top \text{ } \mathbf{iff} \text{ } [\![FearConfirmed\_arise^{i}_{d}\varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top$$

ただし、t は 0 以上の整数とし、上記の iff の右側を満たすような t が複数存在する場合、最小の t をとることにする。また、関数 $f_{FC'}$ は、x,e,l に関する増加関数、関数 f_{FC} は、t に関する減少関数で、$f_{FC}(d,0) = d$ を満たすものとする。関数 f_{FC}、関数 $f_{FC'}$ のどちらも、値域は $[0,1]$ とする。

Relief

Relief の生起条件を以下のように定義する。

$$Relief\_arise^i_{f_{Rl'}(x,e,l)}\ \varphi := Bel^i PFear^i_x \neg\ \varphi\ \wedge\ Bel^i PEffort^i_e \varphi\ \wedge\ Prob\_arise^i_l\ \varphi$$

$Relief^i_{f_{Rl}(d,t)}\varphi$ を新たなオペレータとし、$Relief^i_{f_{Rl}(d,t)}\varphi$ の解釈を以下のように定義する。

$$[\![Relief^i_{f_{Rl}(d,t)}\varphi]\!]_{\langle M,w\rangle} = \top \text{ **iff** } [\![Relief\_arise^i_d \varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top$$

ただし、t は 0 以上の整数とし、上記の iff の右側を満たすような t が複数存在する場合、最小の t をとることにする。また、関数 $f_{Rl'}$ は、x,e,l に関する増加関数、関数 f_{Rl} は、t に関する減少関数で、$f_{Rl}(d,0) = d$ を満たすものとする。関数 f_{Rl}、関数 $f_{Rl'}$ のどちらも、値域は $[0,1]$ とする。

Disappointment

Disappointment の生起条件を以下のように定義する。

$$Disappointment\_arise^i_{f_{Ds'}(x,e,l)}\ \varphi := Bel^i PHope^i_x \neg\ \varphi\ \wedge\ Bel^i PEffort^i_e \varphi\ \wedge\ Prob\_arise^i_l\ \varphi$$

$Disappointment^i_{f_{Ds}(d,t)}\varphi$ を新たなオペレータとし、$Disapointment^i_{f_{Ds}(d,t)}\varphi$ の解釈を以下のように定義する。

$$[\![Disappointment^i_{f_{Ds}(d,t)}\varphi]\!]_{\langle M,w\rangle} = \top \text{ **iff** } [\![Disappointment\_arise^i_d \varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top$$

ただし、t は 0 以上の整数とし、上記の iff の右側を満たすような t が複数存在する場合、最小の t をとることにする。また、関数 $f_{Ds'}$ は、x,e,l に関する増加関数、関数 f_{Ds} は、t に関する減少関数で、$f_{Ds}(d,0) = d$ を満たすものとする。関数 f_{Ds}、関数 $f_{Ds'}$ のどちらも、値域は $[0,1]$ とする。

7.5 Attribution emotions

Pride

Pride の生起条件を以下のように定義する。

$$Pride\_arise^i_{f_{Pr'}(p,l)}\ \varphi := Praise^i_p\ \varphi\ \wedge\ Bel^i PProb^i_l \varphi\ \wedge\ Bel\_arise^i\ \varphi\ \wedge\ l \neq 1$$

$Pride^i_{f_{Pr}(d,t)}\varphi$ を新たなオペレータとし、$Pride^i_{f_{Pr}(d,t)}\varphi$ の解釈を以下のように定義する。

$$[\![Pride^i_{f_{Pr}(d,t)}\varphi]\!]_{\langle M,w\rangle} = \top \text{ **iff** } [\![Pride\_arise^i_d \varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top$$

ただし、t は 0 以上の整数とし、上記の iff の右側を満たすような t が複数存在する場合、最小の t をとることにする。また、関数 $f_{Pr'}$ は、p に関する増加関数、l に関する減少関数であり、関数 f_{Pr} は、t に関する減少関数で、$f_{Pr}(d,0) = d$ を満たすものとする。関数 f_{Pr}、関数 $f_{Pr'}$ のどちらも、値域は $[0,1]$ とする。

Shame

Shame の生起条件を以下のように定義する。

$$Shame\_arise^i_{f_{Sh'}(p,l)}\ \varphi := Praise^i_p\ \varphi \land Bel^i PProb^i_l \varphi \land Bel\_arise^i\ \varphi \land l \neq 1$$

$Shame^i_{f_{Sh}(d,t)}\varphi$ を新たなオペレータとし、$Shame^i_{f_{Sh}(d,t)}\varphi$ の解釈を以下のように定義する。

$$[\![Shame^i_{f_{Sh}(d,t)}\varphi]\!]_{\langle M,w\rangle} = \top \textbf{ iff } [\![Shame\_arise^i_d\varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top$$

ただし、t は 0 以上の整数とし、上記の iff の右側を満たすような t が複数存在する場合、最小の t をとることにする。また、関数 $f_{Sh'}$ は、p,l に関する減少関数であり、関数 f_{Sh} は、t に関する減少関数で、$f_{Sh}(d,0) = d$ を満たすものとする。関数 f_{Sh}、関数 $f_{Sh'}$ のどちらも、値域は $[0,1]$ とする。

Admiration

Admiration の生起条件を以下のように定義する。

$$Admiration\_arise^i_{f_{Ad'}(p,l)}\ \varphi := Bel^i Praise^j_p\ \varphi \land Bel^i PProb^i_l \varphi \land Bel\_arise^i\ \varphi \land l \neq 1$$

$Admiration^i_{f_{Ad}(d,t)}\varphi$ を新たなオペレータとし、$Admiration^i_{f_{Ad}(d,t)}\varphi$ の解釈を以下のように定義する。

$$[\![Admiration^i_{f_{Ad}(d,t)}\varphi]\!]_{\langle M,w\rangle} = \top \textbf{ iff } [\![Admiration\_arise^i_d\varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top$$

ただし、t は 0 以上の整数とし、上記の iff の右側を満たすような t が複数存在する場合、最小の t をとることにする。また、関数 $f_{Ad'}$ は、p に関する増加関数、l に関する減少関数であり、関数 f_{Ad} は、t に関する減少関数で、$f_{Ad}(d,0) = d$ を満たすものとする。関数 f_{Ad}、関数 $f_{Ad'}$ のどちらも、値域は $[0,1]$ とする。

Reproach

Reproach の生起条件を以下のように定義する。

$$Reproach\_arise^i_{f_{Rp'}(p,l)}\ \varphi := Bel^i Praise^j_p\ \varphi \land Bel^i PProb^i_l \varphi \land Bel\_arise^i\ \varphi \land l \neq 1$$

$Reproach^i_{f_{Rp}(d,t)}\varphi$ を新たなオペレータとし、$Reproach^i_{f_{Rp}(d,t)}\varphi$ の解釈を以下のように定義する。

$$[\![Reproach^i_{f_{Rp}(d,t)}\varphi]\!]_{\langle M,w\rangle} = \top \textbf{ iff } [\![Reproach\_arise^i_d\varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top$$

ただし、t は 0 以上の整数とし、上記の iff の右側を満たすような t が複数存在する場合、最小の t をとることにする。また、関数 $f_{Rp'}$ は、p,l に関する減少関数であり、関数 f_{Rp} は、t に関する減少関数で、$f_{Rp}(d,0) = d$ を満たすものとする。関数 f_{Rp}、関数 $f_{Rp'}$ のどちらも、値域は $[0,1]$ とする。

7.6 Well-being/Attribution compounds emotions

Gratification

Gratification の生起条件を以下のように定義する。

$$Gratification\_arise^{i}_{f_{Grtf'}(x,y)}\ \varphi := Joy\_arise^{i}_{x}\ \varphi\ \wedge\ Pride\_arise^{i}_{y}\varphi$$

$Gratification^{i}_{f_{Grtf}(d,t)}\varphi$ を新たなオペレータとし、$Gratification^{i}_{f_{Grtf}(d,t)}\varphi$ の解釈を以下のように定義する。

$$[\![Gratification^{i}_{f_{Grtf}(d,t)}\varphi]\!]_{\langle M,w\rangle} = \top \textbf{ iff } [\![Gratification\_arise^{i}_{d}\varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top$$

ただし、t は 0 以上の整数とし、上記の iff の右側を満たすような t が複数存在する場合、最小の t をとることにする。また、関数 $f_{Grtf'}$ は、x,y に関する増加関数、関数 f_{Grtf} は、t に関する減少関数で、$f_{Grtf}(d,0) = d$ を満たすものとする。関数 f_{Grtf}、関数 $f_{Grtf'}$ のどちらも、値域は $[0,1]$ とする。

Remorse

Remorse の生起条件を以下のように定義する。

$$Remorse\_arise^{i}_{f_{Rm'}(x,y)}\ \varphi := Distress\_arise^{i}_{x}\ \varphi\ \wedge\ Shame\_arise^{i}_{y}\varphi$$

$Remorse^{i}_{f_{Rm}(d,t)}\varphi$ を新たなオペレータとし、$Remorse^{i}_{f_{Rm}(d,t)}\varphi$ の解釈を以下のように定義する。

$$[\![Remorse^{i}_{f_{Rm}(d,t)}\varphi]\!]_{\langle M,w\rangle} = \top \textbf{ iff } [\![Remorse\_arise^{i}_{d}\varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top$$

ただし、t は 0 以上の整数とし、上記の iff の右側を満たすような t が複数存在する場合、最小の t をとることにする。また、関数 $f_{Rm'}$ は、x,y に関する増加関数、関数 f_{Rm} は、t に関する減少関数で、$f_{Rm}(d,0) = d$ を満たすものとする。関数 f_{Rm}、関数 $f_{Rm'}$ のどちらも、値域は $[0,1]$ とする。

Gratitude

Gratitude の生起条件を以下のように定義する。

$$Gratitude\_arise^{i}_{f_{Grtt'}(x,y)}\ \varphi := Joy\_arise^{i}_{x}\ \varphi\ \wedge\ Admiration\_arise^{i}_{y}\varphi$$

$Gratitude^{i}_{f_{Grtt}(d,t)}\varphi$ を新たなオペレータとし、$Gratitude^{i}_{f_{Grtt}(d,t)}\varphi$ の解釈を以下のように定義する。

$$[\![Gratitude^{i}_{f_{Grtt}(d,t)}\varphi]\!]_{\langle M,w\rangle} = \top \textbf{ iff } [\![Gratitude\_arise^{i}_{d}\varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top$$

ただし、t は 0 以上の整数とし、上記の iff の右側を満たすような t が複数存在する場合、最小の t をとることにする。また、関数 $f_{Grtt'}$ は、x,y に関する増加関数、関数 f_{Grtt} は、t に関する減少関数で、$f_{Grtt}(d,0) = d$ を満たすものとする。関数 f_{Grtt}、関数 $f_{Grtt'}$ のどちらも、値域は $[0,1]$ とする。

Anger

Anger の生起条件を以下のように定義する。

$$Anger\_arise^{i}_{f_{An'}(x,y)}\ \varphi := Distress\_arise^{i}_{x}\ \varphi\ \wedge\ Reproach\_arise^{i}_{y}\varphi$$

$Anger^{i}_{f_{An}(d,t)}\varphi$ を新たなオペレータとし、$Anger^{i}_{f_{An}(d,t)}\varphi$ の解釈を以下のように定義する。

$$[\![Anger^{i}_{f_{An}(d,t)}\varphi]\!]_{\langle M,w\rangle} = \top \textbf{ iff } [\![Anger\_arise^{i}_{d}\varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top$$

ただし、t は 0 以上の整数とし、上記の iff の右側を満たすような t が複数存在する場合、最小の t をとることにする。また、関数 $f_{An'}$ は、x,y に関する増加関数、関数 f_{An} は、t に関する減少関数で、$f_{An}(d,0) = d$ を満たすものとする。関数 f_{An}、関数 $f_{An'}$ のどちらも、値域は $[0,1]$ とする。

7.7 Attraction emotions

この感情グループには、時間経過による感情の強さの減衰は考慮しない。というのも、現実世界において好意や嫌悪は時間経過より、イベントに影響を受けると考えられるからである。

Love

OCC theory において、Love の生起条件は、「liking an appealing object」と定義されており、魅力的なオブジェクトに対する好意の感情タイプであり、また、強さに影響する変数については、以下の変数が定義されている。

- オブジェクトがどの程度魅力的か
- オブジェクトとどの程度親しみがあるか

この変数を、魅力の程度を変数に持つオペレータ *Appealing* と、親しみの程度を変数に持つオペレータ *Familiar* で表現し、オペレータ *Love* を以下のように定義する。

$$Love^{i}_{f_L(a,f)}\,j := Appealing(i,j,a)\ \wedge\ Familiar(i,j,f)$$

この関数 f_L は、a,f に関する増加関数とする。

Hate

OCC theory において、Hate の生起条件は、「disliking an unappealing object」と定義されており、非魅力的なオブジェクトに対する嫌悪の感情タイプである。また、強さに影響する変数については、以下の変数が定義されている。

- オブジェクトがどの程度魅力的ではないか

- オブジェクトとどの程度親しみがあるか

この変数を、魅力的ではない程度を変数に持つオペレータ *Unapealing* と、親しみの程度を変数に持つオペレータ *Familiar* で表現し、オペレータ *Hate* を以下のように定義する。

$$Hate^i_{f_{Ha}(u,f)} j := Unappealing(i,j,u) \wedge Familiar(i,j,f)$$

この関数 f_{Ha} は、u, f に関する増加関数とする。

7.8　感情の強さの減衰の表現

ここでは、感情の強さの減衰が本体系でどのように表現されるかを、直感的な議論で例示する。

例えば、ある事象 φ がエージェント i にとって 0.8 ほど望ましく ($Des^i_{0.8}\varphi$)、かつ、ある時刻に φ が起きたことを i が信じたとする。この時、φ が成り立っているという信念はしばらく続くのが自然である。

従来研究では、φ に関する感情 Joy が起きる場合、その強さはその時点で信念 φ があることと、その時の φ の望ましさで決まっていた。望ましさは時刻によって変化しないという仮定を置いているため、信念 φ がある限り、φ に関する Joy の強さは変化しないことになり、感情の強さの減衰は表現できなかった。

これに対して、本論文での体系では、信念 φ が生じた時刻 t に $Bel\_arise^i\varphi$ が成り立ち、この時刻に φ に関する *Joy_arise* がある強さで生起する。以後の時刻において信念 φ が保持されていたとしても、$Bel\_arise^i\varphi$ は成立せず、成り立った時刻 t の *Joy_arise* から次第に減少していく強さで、その時刻での Joy の強さが決まる。

また、このように、信念の保持ではなく生起によって感情の強さが決まるため、信念の生起をトリガリングイベントとしてエージェントの内部状態や行動の変化を引き起こす、実際のエージェントシステムの実装方法とも整合する。

8　感情に関する性質

本章では、感情に関する性質を表す論理式を示す。本論文では、我々の従来研究に比べ、感情の減衰を表現可能となった点が異なる一方で、論理を用いて感情を形式化しているため、従来研究で成り立っていた「感情に関する妥当と思われる性質が、恒真な論理式として示せる」という利点も保持しており、それは本研究で提案した形式化の妥当さを示すものとも言える。

なお、$\vdash \varphi$ は、「φ は恒真である」を表す。

8.1　感情の対立関係

8.1.1　相反する感情

ここで示すのは、同じグループに属する、相反する感情は同時に生起されないという性質である。ここでは、well-being emotions についてのみ説明するが、他の感情グループでも同様に示せるので、他の感情グループについての対立関係については省略する。

- *Joy* の定義に現れる $f_J, f_{J'}$ が、$f_J(f_{J'}(d), 0) = 1.0$ **iff** $d = 1.0$ を満たし、かつ、*Distress* の定義に現れる $f_D, f_{D'}$ が、$f_D(f_{D'}(d), 0) = 1.0$ **iff** $d = 0$ を満たすとき、
  $\vdash \neg(Joy^{i}_{1.0}\varphi \wedge Distress^{i}_{1.0}\varphi)$

- *HappyFor* の定義に現れる $f_{HF}, f_{HF'}$ と、*Distress* の定義に現れる $f_D, f_{D'}$ が、以下のいずれかを満たすとき、

  - $f_{HF}(f_{HF'}(d_i, d_j, v), 0) = 1.0$ **iff** $d_i = 1.0$ かつ、$f_{SF}(f_{SF'}(d_i, d_j, v), 0) = 1.0$ **iff** $d_i = 0$
  - $f_{HF}(f_{HF'}(d_i, d_j, v), 0) = 1.0$ **iff** $d_j = 1.0$ かつ、$f_{SF}(f_{SF'}(d_i, d_j, v), 0) = 1.0$ **iff** $d_j = 0$
  - $f_{HF}(f_{HF'}(d_i, d_j, v), 0) = 1.0$ **iff** $v = 1.0$ かつ、$f_{SF}(f_{SF'}(d_i, d_j, v), 0) = 1.0$ **iff** $v = 0$

  $\vdash \neg(HappyFor^{i,j}_{1.0}\varphi \wedge SorryFor^{i,j}_{1.0}\varphi)$

- *Resentment* の定義に現れる $f_{Rs}, f_{Rs'}$ と、*Gloating* の定義に現れる $f_{Gl}, f_{Gl'}$ が、以下のいずれかを満たすとき、

  - $f_{Rs}(f_{Rs'}(d_i, d_j, v), 0) = 1.0$ **iff** $d_i = 0$ かつ、$f_{Gl}(f_{Gl'}(d_i, d_j, v), 0) = 1.0$ **iff** $d_i = 1.0$
  - $f_{Rs}(f_{Rs'}(d_i, d_j, v), 0) = 1.0$ **iff** $d_j = 1.0$ かつ、$f_{Gl}(f_{Gl'}(d_i, d_j, v), 0) = 1.0$ **iff** $d_j = 0$
  - $f_{Rs}(f_{Rs'}(d_i, d_j, v), 0) = 1.0$ **iff** $v = 0$ かつ、$f_{Gl}(f_{Gl'}(d_i, d_j, v), 0) = 1.0$ **iff** $v = 1.0$

  $\vdash \neg(Resentment^{i,j}_{1.0}\varphi \wedge Gloating^{i,j}_{1.0}\varphi)$

- *Hope* の $f_H, f_{H'}$ が、$f_H(f_{J'}(d), 0) = 1.0$ **iff** $d = 1.0$ を満たし、かつ、*Fear* の $f_F, f_{F'}$ が、$f_F(f_{F'}(d), 0) = 1.0$ **iff** $d = 0$ を満たすとき、
  $\vdash \neg(Hope^{i}_{1.0}\varphi \wedge Fear^{i}_{1.0}\varphi)$

- *Satisfaction* の定義に現れる $f_{St}, f_{St'}$ が、$f_{St}(f_{St'}(x, e, l), 0) = 1.0$ **iff** $x = 1.0$ を満たし、かつ、*FearConfirmed* の定義に現れる $f_{FC}, f_{FC'}$ が、$f_{FC}(f_{FC'}(x, e, l), 0) = 1.0$ **iff** $x = 1.0$ を満たし、かつ、上記の *Hope* と *Fear* に関する条件が満たされるとき、
  $\vdash \neg(Satisfaction^{i}_{1.0}\varphi \wedge FearConfirmed^{i}_{1.0}\varphi)$

- *Relief* の定義に現れる $f_{Rl}, f_{Rl'}$ が、$f_{Rl}(f_{Rl'}(x, e, l), 0) = 1.0$ **iff** $x = 1.0$ を満たし、かつ、*Disappointment* の定義に現れる $f_{Ds}, f_{Ds'}$ が、$f_{Ds}(f_{Ds'}(x, e, l), 0) = 1.0$ **iff** $x = 1.0$ を満たし、かつ、上記の *Hope* と *Fear* に関する条件が満たされるとき、
  $\vdash \neg(Relief^{i}_{1.0}\varphi \wedge Disappointment^{i}_{1.0}\varphi)$

- *Pride* の定義に現れる $f_{Pr}, f_{Pr'}$ が、$f_{Pr}(f_{Pr'}(p, l), 0) = 1.0$ **iff** $p = 1.0$ を満たし、かつ、*Shame* の定義に現れる $f_{Sh}, f_{Sh'}$ が、$f_{Sh}(f_{Sh'}(p, l), 0) = 1.0$ **iff** $p = 0$ を満たすとき、
  $\vdash \neg(Pride^{i}_{1.0}\varphi \wedge Shame^{i}_{1.0}\varphi)$

- *Admiration* の定義に現れる $f_{Ad}, f_{Ad'}$ が、$f_{Ad}(f_{Ad'}(p, l), 0) = 1.0$ **iff** $p = 1.0$ を満たし、かつ、*Reproach* の定義に現れる $f_{Rp}, f_{Rp'}$ が、$f_{Rp}(f_{Rp'}(p, l), 0) = 1.0$ **iff** $p = 0$ を満たすとき、
  $\vdash \neg(Admiration^{i}_{1.0}\varphi \wedge Reproach^{i}_{1.0}\varphi)$

- *Gratification* の定義に現れる $f_{Grtf}, f_{Grtf'}$ と、*Remorse* の定義に現れる $f_{Rm}, f_{Rm'}$ が、以下のいずれかを満たすとき、
  - f_{Grtf} が、$f_{Grtf}(f_{Grtf'}(x,y),0) = 1.0$ **iff** $x = 1.0$ を満たし、かつ、f_{Rm} が、$f_{Rm}(f_{Rm'}(x,y),0) = 1.0$ **iff** $x = 1.0$ かつ、上記の *Joy* と *Distress* に関する条件が満たされる
  - f_{Grtf} が、$f_{Grtf}(f_{Grtf'}(x,y),0) = 1.0$ **iff** $y = 1.0$ を満たし、かつ、f_{Rm} が、$f_{Rm}(f_{Rm'}(x,y),0) = 1.0$ **iff** $y = 1.0$ かつ、上記の *Pride* と *Shame* に関する条件が満たされる

  $\vdash \neg(Gratification^{i}_{1.0}\varphi \wedge Remorse^{i}_{1.0}\varphi)$

- *Gratitude* の定義に現れる $f_{Grtt}, f_{Grtt'}$ と、*Anger* の定義に現れる $f_{An}, f_{An'}$ が、以下のいずれかを満たすとき、
  - f_{Grtt} が、$f_{Grtt}(f_{Grtt'}(x,y),0) = 1.0$ **iff** $x = 1.0$ を満たし、かつ、f_{An} が、$f_{An}(f_{An'}(x,y),0) = 1.0$ **iff** $x = 1.0$ かつ、上記の *Joy* と *Distress* に関する条件が満たされる
  - f_{Grtt} が、$f_{Grtt}(f_{Grtt'}(x,y),0) = 1.0$ **iff** $y = 1.0$ を満たし、かつ、f_{An} が、$f_{An}(f_{An'}(x,y),0) = 1.0$ **iff** $y = 1.0$ かつ、上記の *Admiration* と *Reproach* に関する条件が満たされる

  $\vdash \neg(Gratitude^{i}_{1.0}\varphi \wedge Anger^{i}_{1.0}\varphi)$

- *Love* の定義に現れる f_L が $f_L(a,f) = 1.0$ **iff** $a = 1.0$ を満たし、かつ、*Hate* の定義に現れる f_{Ha} が $f_{Ha}(u,f) = 1.0$ **iff** $u = 1.0$ を満たすとき、

  $\vdash \neg(Love^{i}_{1.0}\varphi \wedge Hate^{i}_{1.0}\varphi)$

8.1.2 友好と嫌悪の感情

友好的な感情と嫌悪的な感情は同時に生起しないという性質がある。これは、エージェントがある事（物）に対して、友好的な感情や嫌悪的な感情を同時に持ち得ないということを証明可能である。

- *HappyFor* の定義に現れる $f_{HF}, f_{HF'}$ と、、*Resentment* の定義に現れる $f_{Rs}, f_{Rs'}$ が、以下のいずれかをを満たすとき、
  - $f_{HF}(f_{HF'}(d_i,d_j,v,LD),0) = 1.0$ **iff** $d_i = 1.0$ かつ、
    $f_{Rs}(f_{Rs'}(d_i,d_j,v,HTD),0) = 1.0$ **iff** $d_i = 0$
  - $f_{HF}(f_{HF'}(d_i,d_j,v,LD),0) = 1.0$ **iff** $v = 1.0$ かつ、
    $f_{Rs}(f_{Rs'}(d_i,d_j,v,HTD),0) = 1.0$ **iff** $v = 0$
  - $f_{HF}(f_{HF'}(d_i,d_j,v,LD),0) = 1.0$ **iff** $LD = 1.0$ かつ、
    $f_{Rs}(f_{Rs'}(d_i,d_j,v,HTD),0) = 1.0$ **iff** $HTD = 1.0$

  $\vdash \neg(HappyFor^{i,j}_{1.0}\varphi \wedge Resentment^{i,j}_{1.0}\varphi)$

- *HappyFor* の定義に現れる $f_{HF}, f_{HF'}$ と、*Gloating* の定義に現れる $f_{Gl}, f_{Gl'}$ が、以下のいずれかを満たすとき、

- $f_{HF}(f_{HF'}(d_i,d_j,v,LD),0)=1.0$ **iff** $d_j=1.0$ かつ、
  $f_{Gl}(f_{Gl'}(d_i,d_j,v,HTD),0)=1.0$ **iff** $d_j=0$
- $f_{HF}(f_{HF'}(d_i,d_j,v,LD),0)=1.0$ **iff** $LD=1.0$ かつ、
  $f_{Gl}(f_{Gl'}(d_i,d_j,v,HTD),0)=1.0$ **iff** $HTD=1.0$

$\vdash \neg(\mathit{HappyFor}^{i,j}_{1.0}\varphi \wedge \mathit{Gloating}^{i,j}_{1.0}\varphi)$

- *SorryFor* の定義に現れる $f_{SF,SF'}$ と、*Resentment* の定義に現れる $f_{Rs,Rs'}$ が、以下のいずれかを満たすとき、
  - $f_{SF}(f_{SF'}(d_i,d_j,v,LD),0)=1.0$ **iff** $d_j=0$ かつ、
    $f_{Rs}(f_{Rs'}(d_i,d_j,v,HTD),0)=1.0$ **iff** $d_j=1.0$
  - $f_{SF}(f_{SF'}(d_i,d_j,v,LD),0)=1.0$ **iff** $LD=1.0$ かつ、
    $f_{Rs}(f_{Rs'}(d_i,d_j,v,HTD),0)=1.0$ **iff** $HTD=1.0$

  $\vdash \neg(\mathit{SorryFor}^{i,j}_{1.0}\varphi \wedge \mathit{Resentment}^{i,j}_{1.0}\varphi)$

- *SorryFor* の定義に現れる $f_{SF,SF'}$ と、*Gloating* の定義に現れる $f_{Gl,Gl'}$ が、以下のいずれかを満たすとき、
  - $f_{SF}(f_{SF'}(d_i,d_j,v,LD),0)=1.0$ **iff** $d_i=1.0$ かつ、
    $f_{Gl}(f_{Gl'}(d_i,d_j,v,HTD),0)=1.0$ **iff** $d_i=1.0$
  - $f_{SF}(f_{SF'}(d_i,d_j,v,LD),0)=1.0$ **iff** $v=1.0$ かつ、
    $f_{Gl}(f_{Gl'}(d_i,d_j,v,HTD),0)=1.0$ **iff** $v=1.0$
  - $f_{SF}(f_{SF'}(d_i,d_j,v,LD),0)=1.0$ **iff** $LD=1.0$ かつ、
    $f_{Gl}(f_{Gl'}(d_i,d_j,v,HTD),0)=1.0$ **iff** $HTD=1.0$

  $\vdash \neg(\mathit{SorryFor}^{i,j}_{1.0}\varphi \wedge \mathit{Gloating}^{i,j}_{1.0}\varphi)$

8.2 予想された事象の実現と非実現

予想された事象が、実現したときに生起される感情と実現しなかったときに生起される感情は同時に持ち得ないという性質がある。

- *Satisfaction* の定義に現れる $f_{St},f_{St'}$ が、$f_{St}(f_{St'}(x,e,l),0)=1.0$ **iff** $l=1.0$ を満たし、かつ、*Disappointment* の定義に現れる $f_{Ds,Ds'}$ が、$f_{Ds}(f_{Ds'}(x,e,l),0)=1.0$ **iff** $l=1.0$ を満たすとき、$\vdash \neg(\mathit{Satisfaction}^{i}_{1.0}\varphi \wedge \mathit{Disappointment}^{i}_{1.0}\neg\varphi)$
- *FearConfirmed* の定義に現れる $f_{FC},f_{FC'}$ が、$f_{FC}(f_{FC'}(x,e,l),t)=1.0$ **iff** $l=1.0$ を満たし、かつ、*Relief* の定義に現れる $f_{Rl},f_{Rl'}$ が、$f_{Rl}(f_{Rl'}(x,e,l),t)=1.0$ **iff** $l=1.0$ を満たすとき、

  $\vdash \neg(\mathit{FearConfirmed}^{i}_{1.0}\varphi \wedge \mathit{Relief}^{i}_{1.0}\neg\varphi)$

8.3　感情の自覚

定義した 22 個の全ての感情 *Emotion* について、以下が成り立つ。

$\vdash Emotion^i_f \varphi \leftrightarrow Bel^i Emotion^i_f \varphi$

$\vdash \neg Emotion^i_f \varphi \leftrightarrow Bel^i \neg Emotion^i_f \varphi$

9　考察

本章では、本論文で提案した論理体系の妥当性と、今後の課題について述べる。

本論文で提案した論理体系によって、時間経過による感情の強さの減衰が表現出きるようになった。それによって、従来では、生起した感情は強さが変わらず、時間が経過してもずっと残ったままになるが、我々の論理体系では、生起した感情の強さが時間の経過によって減衰することが可能となった。これによって、人間の感情のように、時間が経過するにつれて和らいでいく感情を表現することができるようになった。

加えて、信念の生起を考えるようになったこと (p.7 の (1) や (2)) で、より実態および実装を反映出来るようになった。我々の実装では、信念の生起によって感情が生起するため、論理体系と実装の整合性がとれるようになった。

また、Adam らは、感情に関する望ましい性質を恒真な論理式として示すことで、論理体系の妥当性を示している。我々の論理体系は、Adam らの論理体系に準拠し、一階述語論理への拡張したものに新たなオペレータを導入をしているため、Adam らの論理体系で成り立つ感情に関する自然な性質のほとんどは、我々の形式化でもほぼ成り立つ。我々の論理体系では、Adam らの論理体系とは異なり、感情の強さと時間経過による強さの減衰を考慮しなければならないため、感情の強さを決定する関数や変数に制限を加える必要がある。

8.1.1 章で述べた相反する感情について、Adam の論理体系ではいかなるときでも同時に生起することはない。しかし、我々の論理体系では同時刻に対立する感情それぞれの強さが 1.0、すなわち最大であるときのみ、同時に生起されない。それ以外のときには、相反する感情も同時に生起するのだが、相反する感情の強さが最大ではないのであれば、同時に生起することは不自然なことではない。例えば、Adam の論理体系では、$\vdash \neg(Joy^i \varphi \wedge Distress^i \varphi)$ が成り立つが、我々の論理体系では、感情の強さと時間経過による強さの減衰を扱うために、$\vdash \neg(Joy^i_{f_J(1.0,0)} \varphi \wedge Distress^i_{f_D(1.0,0)} \varphi)$ との形で成り立つ。しかし、感情の強さが最大ではなく、同時刻で生起した感情ではないのなら、例えば、$Joy^i_{f_J(0.4,2)} \varphi \wedge Distress^i_{f_D(0.5,1)} \varphi$ が真になることがある。

信念の生起を考えるようになったことで、Adam の論理体系で成り立っていた感情の自覚についての性質が従来研究そのままの形では成り立たなくなった。そのために、時間経過に関する関数に制約 (1P) を加えたが、これは、過去のある時刻に起こったこと、もしくは起こる可能性があったことは、その時刻のどの世界においても変わることはない、すなわち、過去の事実は変わらないということを表しているため、不自然な制約ではない。

また、我々の論理体系では、Adam らの論理体系では一階述語論理を扱えないために省略されていた、好意の感情 (Love) と嫌悪の感情 (Hate) の表現を可能にした。また、OCC theory では Fortunes-of-others Emotions の感情 (例えば、HappyFor など) の生起条件に、相手の好ましさが含まれていたが、Adam らの論理体系では同様に省略されていた。そのた

め、例え赤の他人に対してでも HappyFor が生起するが、これは現実世界において不自然である。我々の論理体系では、一階述語論理を扱えるため、Fortunes-of-others Emotions の生起条件に相手の好ましさを追加した。これによって、Adam らの論理体系で表現可能であった感情より、より人間らしい感情表現が可能となった。

しかし、我々の論理体系で生起可能な感情は自分自身についての感情のみであり、他者の感情の推定については考慮していない。OCC theory で述べられている全ての感情の生起条件をそのまま、他者の感情の推定に用いても妥当かどうかは議論が必要と言える。また、ある事実を知って生起した感情が、その後真逆の事実を知った時に消滅するべきであるが、本論理体系では消滅することはなく残り続けてしまう。

以上の課題に加え、本論文で提案した論理体系の現実世界における妥当性の検討も課題にあげられる。新たに形式化した 2 種の感情も含め、全ての感情について、どの程度人間に近く、もっともであるかの検討が不十分である。また、感情の強さを決定する関数や減衰する関数については、適当な値を返す関数としか決めていない。今回は、簡単化のため、減衰については現在時刻から最も近い時刻に生起した感情から単調に減少していく関数と考えている。しかし、同じ感情を度々生起しているのにもかかわらず、毎回同じ強さの感情を生起するのは、現実世界に即しているかどうかは議論の余地がある。人間の感情は複雑なもので、単調な式で表現可能かという問題もあり、もっともらしさも考慮していかなければならない。

10　終わりに

本論文では、OCC theory で定義されている感情を論理式として形式化した Adam らの論理体系を基に、Adam らの形式化では形式化されていなかった 2 種の感情と、時間経過による感情の強さの減衰を表現可能とするため、新たなオペレータの導入と一階述語論理への拡張することで、従来では表現できなかった 2 種の感情と時間経過による感情の強さの減衰を表現できる論理体系を提案した。

今後は、9 章で述べたような問題を検討していくことが課題となる。

謝辞

本研究の遂行および本論文の作成にあたり、指導教員の新出尚之准教授から丁寧なご指導、助言を賜りました。心からの感謝の気持ちと御礼を申し上げたく、謝辞にかえさせていただきたく思います。

参考文献

[1] Carole Adam, Andreas Herzig, and Dominique Longin. A logical formalization of OCC theory of emotions. Synthese, Vol. 168, No. 2, pp. 201-248, 2009.

[2] Bas R.Steunebrink, Mehdi Dastani, and John-Jules Ch.Meyer. A formal model of emotion triggers:an approach for BDI agents. Synthese, Vol. 185, Issue 1 Supplement, pp. 83-129, 2012.

[3] Rafael H. Bordini, Jomi Fred Hübner, and MichaelWooldridge. Programming MultiAgent Systems in AgentSpeak using Jason. John Wiley & Sons, 2007.

[4] A. Ortony, G. L. Clore, and A. Collins. The Cognitive Structure of Emotions. Cambridge University Press, 1988.

[5] Anand S. Rao, Munindar P. Singh, and Michael P. Georgeff. Formal Methods in DAI:Logic-Based Representation and Reasoning. Massachusetts Institute of Technology, 1999.

[6] 藤田恵, 片山寛子, 新出尚之, 高田司郎. 実世界の多様性に適応した BDI ロボットについて. 情報処理学会論文誌数理モデル化と応用, Vol. 5, No. 1, pp. 50-64, 2012.

[7] 山根瑞樹. OCC theory による感情表現を持つエージェントの実現. 2013 年度卒業論文, 奈良女子大学理学部情報科学科, 2014.

[8] 石川葉子. 感情に程度の強さを取り入れたエージェントの実現. 2014 年度卒業論文. 奈良女子大学理学部情報科学科, 2015.

[9] 池之内彰子. OCC theory に基づくエージェントの感情表現の論理モデル. 2014 年度卒業論文, 奈良女子大学人間文化研究科情報科学専攻, 2015.

[10] 向井香里. 時間経過による感情の程度の減衰を取り入れたエージェントの実現. 2015 年度卒業論文, 奈良女子大学理学部情報科学科, 2016.

[11] 今井那緒. 複数の感情の同時生起および削除を取り入れたエージェントの実装. 2015 年度卒業論文, 奈良女子大学理学部情報科学科, 2016.

[12] 浅井沙良. エージェントの感情生起についてー新たな感情の追加ー. 2017 年度卒業論文, 奈良女子大学生活環境学部情報衣環境学科生活情報通信科学コース, 2018(発表予定).

[13] 塚本麻衣. 感情表現を持つ自律エージェント. 2017 年度卒業論文, 奈良女子大学生活環境学部情報衣環境学科生活情報通信科学コース, 2018(発表予定).

付録

ここでは、6 章と 8 章で記述した、本研究での新たに導入したオペレータに関する恒真論理式と、感情に関する性質を表す恒真論理式の証明を示す。

(D-*Bel_arise*)

$$Bel\_arise^i\varphi \to \neg Bel\_arise^i\neg\varphi$$

証明

ある世界 w において、$[\![Bel\_arise^i\varphi \to \neg Bel\_arise^i\neg\varphi]\!]_{\langle M,w\rangle} = \bot$ と仮定する。

$$[\![Bel\_arise^i\varphi]\!]_{\langle M,w\rangle} = \top \tag{3}$$

$$[\![\neg Bel\_arise^i\neg\varphi]\!]_{\langle M,w\rangle} = \bot \tag{4}$$

(3) より、

$$[\![\neg YBel^i\varphi]\!]_{\langle M,w\rangle} = \top \tag{5}$$

$$[\![Bel^i\varphi]\!]_{\langle M,w\rangle} = \top \tag{6}$$

また、(4) より、$[\![Bel\_arise^i\neg\varphi]\!]_{\langle M,w\rangle} = \top$ である。よって、$[\![\neg YBel^i\neg\varphi \wedge Bel^i\neg\varphi]\!]_{\langle M,w\rangle} = \top$ であるため、

$$[\![\neg YBel^i\neg\varphi]\!]_{\langle M,w\rangle} = \top \tag{7}$$

$$[\![Bel^i\neg\varphi]\!]_{\langle M,w\rangle} = \top \tag{8}$$

ここで (D-Bel^i) より、(6) と (8) は矛盾する。
従って、$Bel\_arise^i\varphi \to \neg Bel\_arise^i\neg\varphi$ は恒真である。

(4-*Bel_arise*^{i})

$$Bel\_arise^i\varphi \to Bel^iBel\_arise^i\varphi$$

証明

ある世界 w において、$[\![Bel\_arise^i\varphi \to Bel^iBel\_arise^i\varphi]\!]_{\langle M,w\rangle} = \bot$ と仮定すると、

$$[\![Bel\_arise^i\varphi]\!]_{\langle M,w\rangle} = \top \tag{9}$$

$$[\![Bel^iBel\_arise^i\varphi]\!]_{\langle M,w\rangle} = \bot \tag{10}$$

となる。(9) より、$[\![\neg YBel^i\varphi \wedge Bel^i\varphi]\!]_{\langle M,w\rangle} = \top$ となり、

$$[\![YBel^i\varphi]\!]_{\langle M,w\rangle} = \bot \tag{11}$$

$$[\![Bel^i\varphi]\!]_{\langle M,w\rangle} = \top \tag{12}$$

となる。$w_1 = \mathscr{X}^{-1}(w)$ をみたす w_1 が存在して、(11) より、

$$[\![Bel^i\ \varphi]\!]_{\langle M,w_1\rangle} = \bot \tag{13}$$

となる。

また、(10) より、$[\![Bel^i(\neg YBel^i\varphi \wedge Bel^i\varphi)]\!]_{\langle M,w\rangle} = \bot$ となる。$(C-\Box)$ より、$[\![Bel^i\neg YBel^i\ \varphi \wedge Bel^iBel^i\ \varphi]\!]_{\langle M,w\rangle} = \bot$ である。ここで、$(4-Bel^i)$ と (12) より、$[\![Bel^iBel^i\ \varphi]\!]_{\langle M,w\rangle} = \top$ であるから、$[\![Bel^i\neg YBel^i\ \varphi]\!]_{\langle M,w\rangle} = \bot$ である。

$\mathscr{B}^i(w,w_2) = d > 0$ を満たすある世界 w_2 が存在して、その世界 w_2 において、$[\![\neg YBel^i\ \varphi]\!]_{\langle M,w_2\rangle} = \bot$ である。すなわち、$[\![YBel^i\ \varphi]\!]_{\langle M,w_2\rangle} = \top$ である。

$w_3 = \mathscr{X}^{-1}(w_2)$ を満たすある世界 w_3 が存在して、その世界 w_3 において、

$$[\![Bel^i\ \varphi]\!]_{\langle M,w_3\rangle} = \top \tag{14}$$

ここで、$(1P)$ より、(13) と (14) は矛盾する。

従って、$Bel\_arise^i\ \varphi \to Bel^iBel\_arise^i\ \varphi$ は、恒真である。

$Bel^iBel\_arise^i\ \varphi \to Bel\_arise^i\ \varphi$

また、$Bel^iBel\_arise^i\ \varphi \to Bel\_arise^i\ \varphi$ も恒真である。

証明

ある世界 w において、$[\![Bel^iBel\_arise^i\ \varphi \to Bel\_arise^i\ \varphi]\!]_{\langle M,w\rangle} = \bot$ と仮定すると、

$$[\![Bel^iBel\_arise^i\ \varphi]\!]_{\langle M,w\rangle} = \top \tag{15}$$

$$[\![Bel\_arise^i\ \varphi]\!]_{\langle M,w\rangle} = \bot \tag{16}$$

となる。(15) より、$\mathscr{B}^i(w,w_1) = d > 0$ を満たす全ての世界 w_1 において、$[\![Bel\_arise^i\ \varphi]\!]_{\langle M,w_1\rangle} = \top$ となる。よって、$[\![\neg YBel^i\ \varphi \wedge Bel^i\ \varphi]\!]_{\langle M,w_1\rangle} = \top$ であるから、

$$[\![YBel^i\ \varphi]\!]_{\langle M,w_1\rangle} = \bot \tag{17}$$

$$[\![Bel^i\ \varphi]\!]_{\langle M,w_1\rangle} = \top \tag{18}$$

となる。

また、(16) より、$[\![\neg YBel^i \wedge Bel^i\ \varphi]\!]_{\langle M,w\rangle} = \bot$ である。ここで、

$$[\![Bel^i\ \varphi]\!]_{\langle M,w\rangle} = \top \tag{19}$$

$$[\![\neg YBel^i\ \varphi]\!]_{\langle M,w\rangle} = \bot \tag{20}$$

と

$$[\![Bel^i\ \varphi]\!]_{\langle M,w\rangle} = \bot \tag{21}$$

$$[\![\neg YBel^i\ \varphi]\!]_{\langle M,w\rangle} = \top \tag{22}$$

の二つに場合分けできる。

(20) より、$[\![YBel^i\ \varphi]\!]_{\langle M,w\rangle} = \top$ となり、これと (17) が $(1P)$ より矛盾する。

(21) と (18) が $(4-Bel^i)$ より矛盾する。

以上より、$Bel^iBel\_arise^i\ \varphi \to Bel\_arise^i\ \varphi$ は、恒真である。

よって、$Bel\_arise^i\varphi \leftrightarrow Bel^iBel\_arise^i\varphi$ も恒真であることが証明できる。

(5-*Bel_arise*^{i})

$$\neg Bel\_arise^i\varphi \to Bel^i\neg Bel\_arise^i\varphi$$

証明

ある世界 w において、$[\![\neg Bel\_arise^i\varphi \to Bel^i\neg Bel\_arise^i\varphi]\!]_{\langle M,w\rangle} = \bot$ と仮定する。

$$[\![\neg Bel\_arise^i\varphi]\!]_{\langle M,w\rangle} = \top \tag{23}$$

$$[\![Bel^i\neg Bel\_arise^i\varphi]\!]_{\langle M,w\rangle} = \bot \tag{24}$$

(23) より、$[\![\neg YBel^i\varphi \wedge Bel^i\varphi]\!]_{\langle M,w\rangle} = \bot$ である。また、(24) より、$\mathscr{B}^i(w,w_1) > 0$ を満たす、ある世界 w_1 において、$[\![\neg Bel\_arise^i\varphi]\!]_{\langle M,w_1\rangle} = \bot$ である。

$$[\![Bel\_arise^i\varphi]\!]_{\langle M,w_1\rangle} = \top \tag{25}$$

(5-*Bel*^{i}) より、$[\![Bel^i\varphi]\!]_{\langle M,w\rangle} = \top$ である。従って、(23) より、$[\![\neg YBel^i\varphi]\!]_{\langle M,w\rangle} = \bot$。すなわち、$w_2 = \mathscr{X}^{-1}(w)$ を満たす世界 w_2 において、

$$[\![Bel^i\varphi]\!]_{\langle M,w_2\rangle} = \top \tag{26}$$

(25) より、$[\![\neg YBel^i\varphi \wedge Bel^i\varphi]\!]_{\langle M,w_1\rangle} = \top$ である。よって、

$$[\![Bel^i\varphi]\!]_{\langle M,w_1\rangle} = \top \tag{27}$$

$$[\![\neg YBel^i\varphi]\!]_{\langle M,w_1\rangle} = \top \tag{28}$$

(28) より、$w_3 = \mathscr{X}^{-1}(w_1)$ を満たす世界 w_3 において、$[\![Bel^i\varphi]\!]_{\langle M,w_3\rangle} = \bot$ である。これと、(26) より、(1P) に矛盾する。

従って、$\neg Bel\_arise^i\varphi \to Bel^i\neg Bel\_arise^i\varphi$ は、恒真である。

8.3 感情の自覚

$$\vdash Emotion^i_{f(d,t)}\ \varphi \leftrightarrow Bel^iEmotion^i_{f(d,t)}\ \varphi$$

全ての感情についての証明は省略し、Joy について証明する。その他の感情についても同様に証明できる。

証明

$Joy^i_{f_J(d,t)}\ \varphi \leftrightarrow Bel^i Joy^i_{f_J(d,t)}\ \varphi$ を示す。

ある世界 w において、$[\![Joy^i_{f_J(d,t)}\ \varphi]\!]_{\langle M,w\rangle} = \top$ より、Joy の生起条件から、ある d,t が存在して、

$$[\![Joy\_arise^i_d\ \varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top$$

よって、$d = f_{J'}(d_1)$ を満たす、ある d_1 が存在して、$[\![Des^i_{d_1}\ \varphi \wedge Bel\_arise^i\ \varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top$ であるから、

$$[\![Des^i_{d_1}\ \varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top \tag{29}$$

$$[\![Bel\_arise^i\ \varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top \tag{30}$$

(4-MIX) と (29)、(4-Bel_arise^i) と (30) より、

$$[\![Bel^i Des^i_{d_1}\ \varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top \tag{31}$$

$$[\![Bel^i Bel\_arise^i\ \varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top \tag{32}$$

w から $\mathscr{B}^i$ で到達可能な世界 w_1 において、(31) と (32) と (1P) より、

$$[\![Bel^i Des^i_{d_1}\ \varphi]\!]_{\langle M,\mathscr{X}^{-t}(w_1)\rangle} = \top \tag{33}$$

$$[\![Bel^i Bel\_arise^i\ \varphi]\!]_{\langle M,\mathscr{X}^{-t}(w_1)\rangle} = \top \tag{34}$$

(4-MIX) と (35)、(4-Bel_arise^i) と (36) より、

$$[\![Des^i_{d_1}\ \varphi]\!]_{\langle M,\mathscr{X}^{-t}(w_1)\rangle} = \top \tag{35}$$

$$[\![Bel\_arise^i\ \varphi]\!]_{\langle M,\mathscr{X}^{-t}(w_1)\rangle} = \top \tag{36}$$

従って、$[\![Des^i_{d_1}\ \varphi \wedge Bel\_arise^i\ \varphi]\!]_{\langle M,\mathscr{X}^{-t}(w_1)\rangle} = \top$ であるから、$[\![Joy\_arise^i_d\ \varphi]\!]_{\langle M,\mathscr{X}^{-t}(w_1)\rangle} = \top$、すなわち、$[\![Joy^i_{f_J(d,t)}\ \varphi]\!]_{\langle M,w_1\rangle} = \top$。よって、$\mathscr{B}^i(w,w_1) > 0$ であるから、

$$[\![Bel^i Joy^i_{f_J(d,t)}\ \varphi]\!]_{\langle M,w\rangle} = \top$$

従って、 $Joy^i_{f_J(d,t)}\ \varphi \rightarrow Bel^i Joy^i_{f_J(d,t)}\ \varphi$ は、恒真である。

また、ある世界 w において、$[\![Bel^i Joy^i_{f_J(d,t)}\ \varphi]\!]_{\langle M,w\rangle} = \top$ より、Joy の生起条件から、ある d,t が存在して、

$$[\![Bel^i Joy\_arise^i_d\ \varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top$$

よって、$d = f_{J'}(d_1)$ を満たす、ある d_1 が存在して、$[\![Bel^i(Des^i_{d_1}\ \varphi \wedge Bel\_arise^i\ \varphi)]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top$ である。また、(C-□) より、$[\![Bel^i Des^i_{d_1}\ \varphi \wedge Bel^i Bel\_arise^i\ \varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top$ であるから、

$$[\![Bel^i Des^i_{d_1}\ \varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top \tag{37}$$

$$[\![Bel^i Bel\_arise^i\ \varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top \tag{38}$$

(37) と (4-MIX)、(38) と (4-Bel_arise^i) より、

$$[\![Des^i_{d_1}\ \varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top \tag{39}$$

$$[\![Bel\_arise^i\ \varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top \tag{40}$$

これらより、$[\![Des^i_{d_1}\ \varphi \wedge Bel\_arise^i\ \varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top$ となり、$[\![Joy\_arise^i_d\ \varphi]\!]_{\langle M,\mathscr{X}^{-t}(w)\rangle} = \top$ となるため、

$$[\![Joy^i_{f_J(d,t)}\ \varphi]\!]_{\langle M,w\rangle} = \top$$

従って、 $Bel^i Joy^i_{f_J(d,t)}\ \varphi \rightarrow Joy^i_{f_J(d,t)}\ \varphi$ は、恒真である。

以上より、$Bel^i Joy^i_{f_J(d,t)}\ \varphi \leftrightarrow Joy^i_{f_J(d,t)}\ \varphi$ は、恒真である。