

# OCC Theory に基づくエージェントの感情表現の論理モデル

奈良女子大学 人間文化研究科 情報科学専攻

13840030 池之内彰子

2015年1月30日

## 概要

近年、エージェントと人との対話を実現させる上で、感情を持ち、その感情を行動決定に活かせるといった、より人間に近いエージェントが求められている。OCC Theory と呼ばれる、心理学的見地に基づいて、22種類の感情を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た研究がある。しかし、これらの研究では感情の強さについては扱っておらず、感情は「生起される」または「生起されない」という状態しか扱えないという課題がある。そこで本論文では、基本的に、従来の研究による形式化を踏まえつつ、OCC Theory で定義されている、感情の強さに影響する変数を導入することにより、度合い付きの感情を表現できる論理体系を提案する。

# 1 はじめに

近年、感情を持つエージェントが、研究者の大きな関心の的となっている。そして、人間とのコミュニケーションを重要視するロボットが多く出現し、人間に似た外見を持つことよりも、感情を持ち、その感情をふるまいやしぐさなどの行動に活かせるといった、より人間的なエージェントが求められている。特に、人間の生活に溶け込み、コミュニケーションを必要とするようなエージェントは、実世界において、感情を持ちつつ行動できることが望ましい。これらの背景から、我々は、感情を生起し、自律的に行動決定ができる BDI エージェントの実現を目指している。

Ortony、Clore、Collins らによる、心理学的見地を基にした 22 種類の感情のタイプをモデル化した理論 (以下 OCC Theory と呼ぶ)[1] は、人間の包括的な感情を形式化しており、信念や願望などの心的状態を用いた感情の特徴付けが明確である。そして、その感情の特徴付けを論理モデルとして表現可能な形で行える。一方、BDI エージェントの基盤となる、BDI モデルとは、信念、願望、意図の 3 つの心的パラメータを明示的に持ち、これらの心的パラメータを保持・更新することで意思決定を行い、目標を達成するように振る舞う、人間の合理的な行動をモデル化したもので、BDI logic という論理モデルを持つ。

BDI モデルと OCC Theory のどちらも、論理モデルを利用であるため、親和性が高いと言える。そして、この BDI モデルと OCC Theory を用いてエージェントの感情を扱っている関連研究としては、Adam らによる [4] や、Steunebrink らによる [5] がある。

Adam ら [4] は、BDI モデルの形式的記述のための論理体系である BDI logic に対し、独自に新たなオペレータを導入して拡張し、OCC Theory で扱われている感情の定義を論理式として形式化することで、BDI モデルに取り込んでいる。そこで我々は、これまでに、Adam らの形式化を踏まえ、感情を生起し、その感情によって行動の選択が可能なエージェントの実装を試み、BDI アーキテクチャに基づくエージェント記述言語 AgentSpeak の処理系である Jason[11] で実現、ライブラリを実装した [7][8][9]。

また、Steunebrink ら [5] は、Adam らの形式化には、多くの仮定が含まれており、OCC Theory で定義されている感情の生起条件というよりは、感情の状態を表す形式化であると述べている。そして、Adam らの形式化よりさらに、OCC Theory に忠実になるよう、定義されている感情の生起条件を基に多くのオペレータを導入し、22 種類全ての感情を論理式として表現できる論理体系を提案している。

しかし、Adam らの研究や Steunebrink らの研究における形式化では、OCC Theory で定義されている、感情の生起条件に焦点を当てており、感情の強さに影響を与える変数については考慮されていない。つまり、感情は「生起される」または「生起されない」という状態しか扱えず、生起された感情の強さについては扱えない。また、一度生起した感情は時間が経過しても消滅しないという課題がある。

これらの研究の他には、OCC Theory を用いず、独自の感情の概念を取り入れ、BDI エージェントを構築した、Pereira らによる [6] があり、特に恐怖の感情に焦点を当て形式化し、恐怖の感情に対して、強度についても扱っている。この研究では、CTL の時相オペレータを用いて、恐怖の強さを 3 段階に分け、それぞれ別の論理式として定義している。しかしこの方法では、ひとつの感情に対し、感情の強さはあらかじめ決められた、3 段階の強さしか定義できない問題がある。また、この研究では、恐怖の感情のみを扱っており、他の感情については扱っていない。

本論文では、基本的に Adam らの形式化を踏まえつつ、OCC Theory で定義されている、感情の強さに影響を与える変数を導入することで、より人間に近いエージェントを実現するための、度合い付きの感情を表現できる論理体系を提案する。なお、本研究では、生起された感情の強さのみ焦点を当てることにし、時間経過による感情の消滅については扱わない。

## 2 OCC Theory

本章では、本研究と、先行研究である [4] や [5] で用いられた、心理学的見地に基づき 22 種類の感情を形式化した理論、OCC Theory[1] について詳しく述べる。

OCC Theory は、その簡潔さと実現性が評価され、心理学の分野だけでなく、計算機科学の分野でも用いられており、感情を持つエージェントの構成に広く利用されている。その理由は、人間の感情を 22 種類の感情タイプとして包括的に形式化しており、感情の生起条件が明確で、また、感情の強さに関する評価変数が有限個で定義されており、エージェントの構成には十分であると考えられるためである。

Ortony らは、感情の定義を「感情とは、人が心的過程の中で行うさまざまな情報処理のうちで、人、物、出来事、環境についてする評価的な反応である」と記述している。そして大平 [12] は、この「評価」というのは、対象を、良い - 悪い、安全 - 危険、有用 - 有害、好き - 嫌い、などの軸に位置づけ、認識することであり、また、「反応」とは、対象による脳、神経、身体器官の作用から、潜在的な行動の準備状態の形成、顕在的な表情や行動の表出、主観的な心的体験まで、広い範囲を含むと説明している。

Ortony らは、この定義に従い、22 種類の感情のタイプを、反応の対象で大きく 3 つのクラス “Reaction to events”、“Reaction to agents”、“Reaction to objects” に分けており、さらに、この 3 つのクラスは、いくつかのグループに細分化され、22 種類の感情タイプはそれぞれのグループに属することになる。この感情タイプの分類については、OCC Theory([1], p.19) に、図 1 のようにまとめられている。

まず、“Reaction to events” クラスは、起こったイベントの結果が自分自身にとって望ましいかどうかという点に焦点を当てた、2 個の感情タイプ (Joy, Distress) から成る ‘Well-being Emotions’ グループ、起こったイベントの結果が他者にとって望ましいかどうか、またそのイベントの結果が自分にとって望ましいかどうか、という点に焦点を当てた、4 個の感情タイプ (Happy-For, Sorry-For, Resentment, Gloating) から成る ‘Fortunes-of-others Emotions’ グループ、予想したイベントが望ましいかどうかに焦点を当てた感情タイプ (Hope, Fear)、予想したイベントが起こった結果が望ましいかどうかに焦点を当てた感情タイプ (Satisfaction, FearConfirmed, Relief, Disappointment)、以上 6 個の感情タイプから成る ‘Prospect-based Emotions’ グループ、以上の 3 つのグループに分けられる。

次に、“Reaction to agents” クラスには、イベントの結果ではなく、自分もしくは他者の行動に対する評価に焦点を当てた、4 個の感情タイプ (Pride, Shame, Admiration, Reproach) からなる ‘Attribution Emotions’ グループがある。

また、“Reaction to objects” クラスには、対象の魅力に対する評価に焦点を当てた、2 個の感情タイプ (Love, Hate) から成る ‘Attraction Emotions’ グループがある。

最後に、‘Well-being Emotions’ グループと ‘Attribution Emotions’ グループの複合的なグループがあり、自分もしくは他者の行動と、イベントの結果が望ましいかどうかの両方に焦点を当てた、4 個の感情タイプ (Gratification, Remorse, Gratitude, Anger) からなる ‘Compounds Emotions’ グループがある。

加えて、“Reaction to events” クラスは、イベントの結果が望ましいか、望ましくないかという評価、“Reaction to agents” クラスは、エージェントの行動が称賛されるか、称賛されないかという評価、“Reaction to objects” クラスは、対象が好みか、好みでないかという評価が主要変数として定義されている (図 2([1], p.69))。

OCC Theory で定義されている 22 種類の感情タイプには、感情タイプの指定として、事象に対するリアクションと生起条件、そのタイプに分類される感情の例、感情の強さに影響する変数がそれぞれ定義されている。同じグループに属する感情タイプは、生起条件が同じ形で定義されて

おり、感情の強さに影響する変数についても、同じ変数が用いられている。ここで、変数については、パラメータと呼ぶ方が妥当ではあるが、OCC Theory では変数という単語を使用しているため、本論文でも変数と呼ぶこととする。

Adam ら [4] はこの OCC Theory と Lazarus[2] のバーチャルエージェントの感情表現に適した理論等と比較した上で、OCC Theory は精密で明確な理論であり、OCC Theory を採用する利点として、以下の4つの利点を挙げている。(我々が本研究の出発点とした Adam らの研究については次章で詳しく記述する。)

- OCC Theory は非常に簡潔で実現性が高く、計算機科学者の期待と要求に一致する。OCC Theory が有限の評価変数の組み合わせを用いていることは、現状のアプリケーションにとっては十分であり、感情を持つエージェントの構成に幅広く利用されている。
- 感情には良し悪しの評価が伴うことが必要で、その感情につながる評価はいつも同じであることが要求されると OCC Theory に記述されている。つまり、この良し悪しに関する評価を伴うものが感情である。それゆえ、surprise(驚き) や being abandoned(自暴自棄) などのような、良し悪しの評価が一定でない状態は感情から排除される。さらに、良し悪しの評価は論理で自然に捉えることができるため、OCC Theory は論理的な形式化に非常に適合する。
- OCC Theory は、簡潔で明確な木構造を持ち、また、信念、願望、基準のような、論理学の分野でよく研究された概念を用いており、このことは、形式化の作業を容易にする。
- OCC Theory は包括的な理論であり、このことは、堅牢で多目的なエージェントの構築に重要である。つまり、多種多様なシチュエーションに、感情的に反応できるエージェントが構築できるということである。

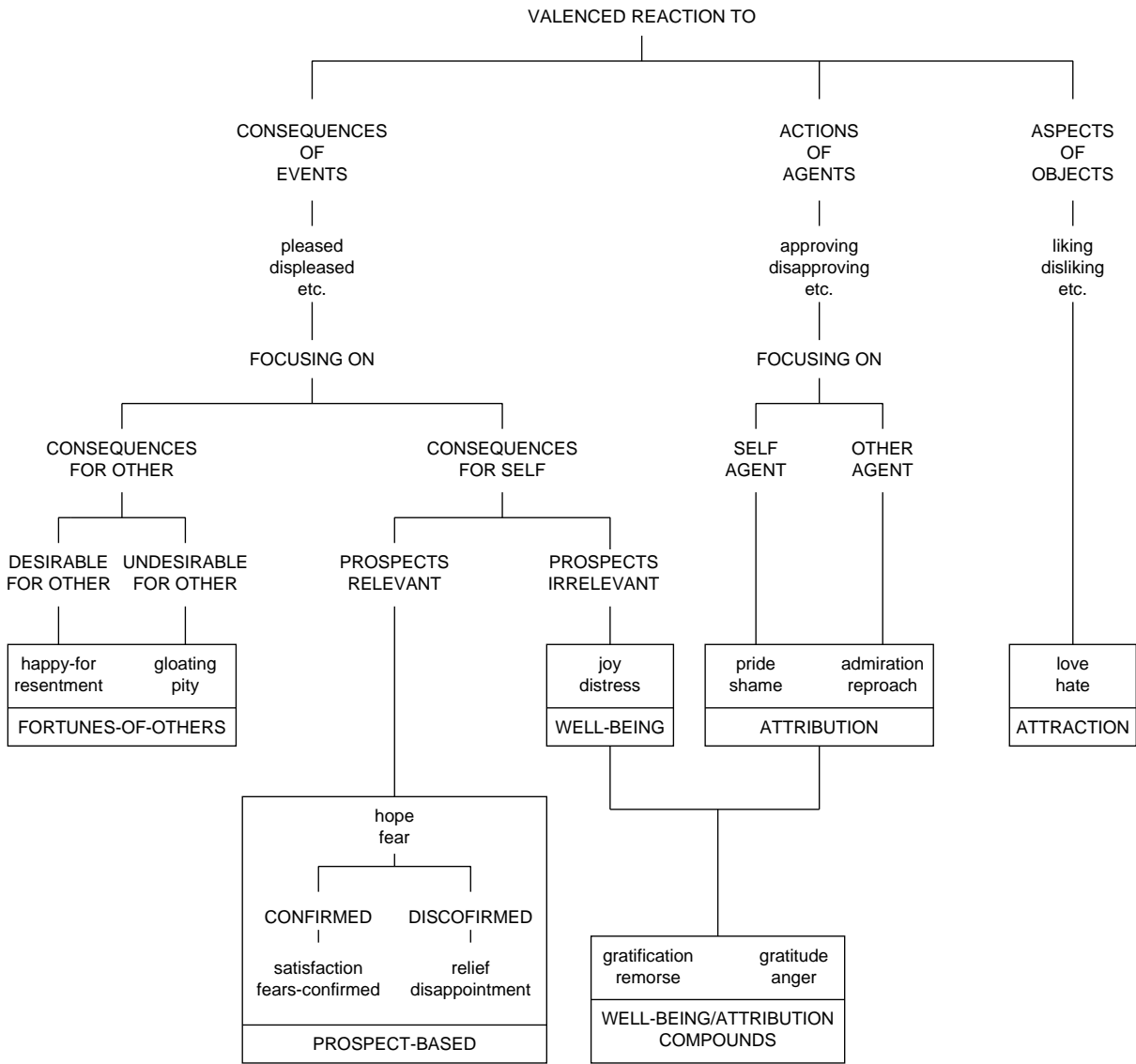


図 1: 感情タイプの分類

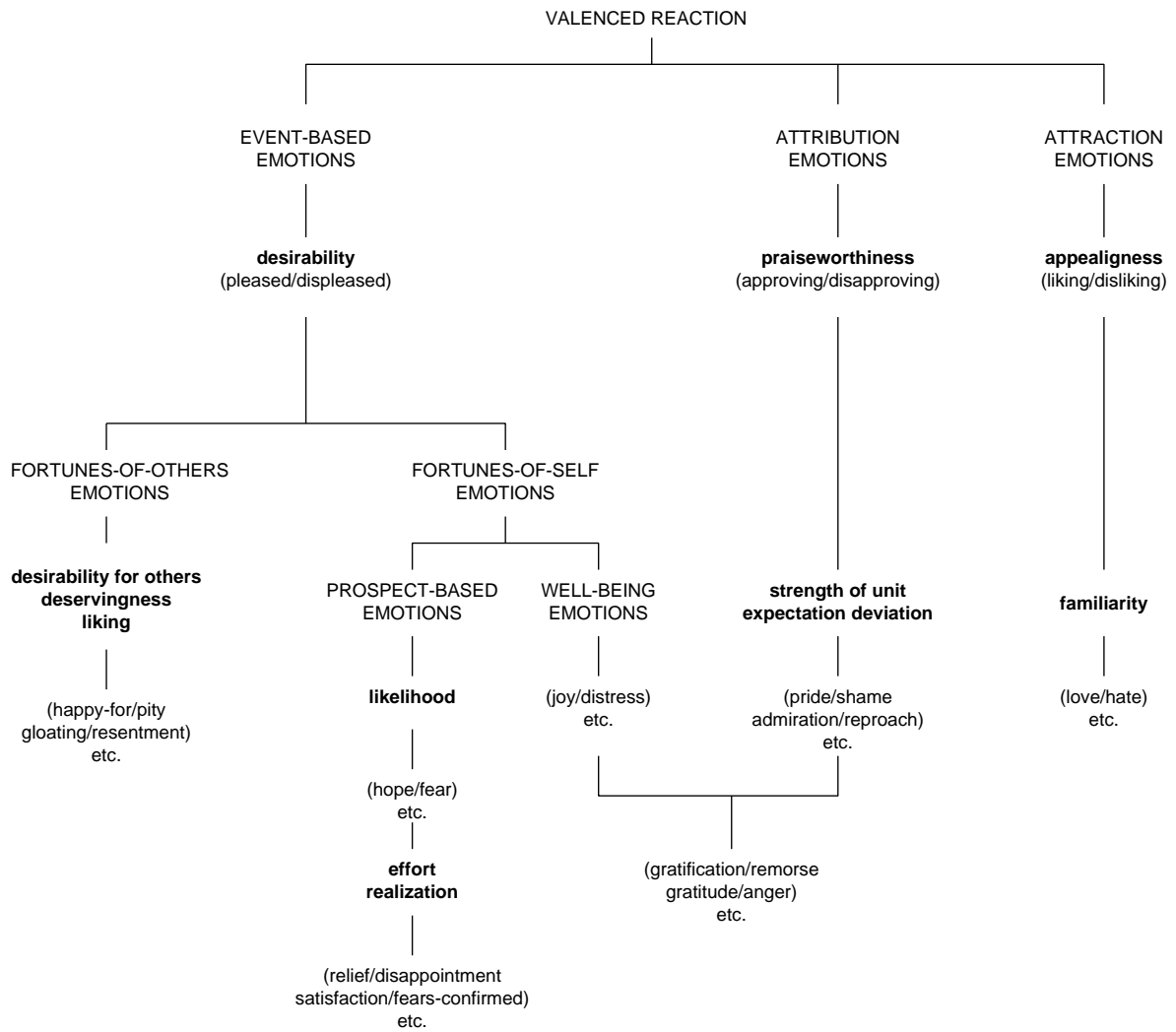


図 2: 強さに影響するローカル変数の分類

## 3 先行研究

### 3.1 Adam らによる形式化

Adam ら [4] 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。そしてその BDI モデルの形式的記述のための論理体系である BDI logic に対し、独自の新たなオペレータを導入して拡張し、OCC Theory で扱われる感情の定義を論理式として形式化することで、BDI モデルに取り込んでいる。

しかし、Adam らの形式化では、Attraction Emotions グループに属する Love と Hate の感情タイプが扱えない。これは、対象物を引数としてとる必要があるため、一階述語論理が必要になり、Adam らの論理体系では、述語論理が扱えないため、形式化を省いている。

また、この研究における形式化では、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。つまり、感情は「生起されている」もしくは「生起されていない」という状態しか扱えず、生起された感情の強さについては扱えないという課題がある。

上述した課題はあるものの、Adam らの形式化は、感情に関する専門家である心理学者の理論 OCC Theory と、論理学と計算モデルの特長を融合した、非常に学際的かつ斬新な研究である。エージェントの感情表現に BDI モデルを用いるような研究の先頭に位置する研究であるため、我々は本研究の出発点として Adam らの研究を選択した。

### 3.2 Steunebrink らによる形式化

Steunebrink ら [5] は、Adam らの形式化には、多くの仮定が含まれており、感情の生起条件というよりは、感情の状態を表す形式化であると述べ、Adam らの形式化よりさらに、OCC Theory に忠実になるよう、定義されている感情の生起条件と、感情の強さに影響する変数を基に多くのオペレータを導入している。特に、感情を引き起こすトリガーを強調しており、22 種類の感情を論理式として表現できる論理体系を提案している。しかし、Steunebrink らの形式化も OCC Theory の感情の強さに影響する評価変数を完全に網羅しているわけではない。

そして、Adam らの研究と同様、生起された感情の強さについては扱えないということと、オペレータの種類が非常に多く、直感的な理解が困難であるという問題点がある。また、前述したように、Steunebrink らは感情のトリガーに焦点を当てているが、我々は、感情が既に生起されているかどうかには焦点を当てているので、Adam らの形式化を採用することにした。

### 3.3 その他の形式化

Adam らや Steunebrink らの他には、Pereira らによる [6] がある。この研究は、OCC Theory を用いず、独自の感情の概念を取り入れ、BDI エージェントを構築しており、特に恐怖の感情に焦点を当て形式化し、それらの感情に対しては、強度についても扱っている。この研究では、CTL の時相オペレータを用いて、恐怖を引き起こすと考えられる要因 *threats* と *unpleasant* の 2 つを 3 段階に分け、それぞれ別の論理式として定義している。例えば、*threats* は Dangerous Threat, Serious Threat, Possible Threat のように分け、*unpleasant* は、Highly Unpleasant, Strongly Unpleasant, Possibly Unpleasant のように分けている。しかしこの方法では、感情の強さは、あらかじめ決められた、感情を引き起こす要因につき 3 段階しか定義できない問題がある。また、この研究では、恐怖の感情のみを扱っており、他の感情については扱っていない。

## 4 構文

本章では、論理式の定義を与える。

- 命題記号の有限集合を  $ATM = \{P_1, P_2, \dots, P_n\}$  とする。命題記号は論理式である。
- エージェントを表す記号の有限集合を  $AGT = \{i_1, i_2, \dots, i_n\}$  とする。
- $\phi, \psi$  が論理式ならば、 $\neg\phi, \phi \vee \psi$  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。
  - $\phi \wedge \psi := \neg(\neg\phi \vee \neg\psi)$
  - $\phi \rightarrow \psi := \neg\phi \vee \psi$
  - $\phi \leftrightarrow \psi := (\phi \rightarrow \psi) \wedge (\psi \rightarrow \phi)$
- $\phi$  が論理式、 $i \in AGT$ 、 $d$  (desirability) が  $0 \leq d \leq 1$  を満たす実数のとき、 $Des_d^i \phi$  は論理式であるとする。  
直感的な意味は、「エージェント  $i$  にとって  $\phi$  は程度  $d$  で望ましい」となる。
- $\phi$  が論理式、 $i \in AGT$ 、 $\ell$  (likelihood) が  $0 \leq \ell \leq 1$  を満たす実数のとき、 $Prob_\ell^i \phi$  は論理式であるとする。また省略形として、以下を定義する。

$$Bel^i \phi := Prob_{1,0}^i \phi$$

直感的な意味は、 $Prob_\ell^i \phi$  が「エージェント  $i$  は、見込み  $\ell$  で、 $\phi$  が真になる可能性がある」と信じている、 $Bel^i \phi$  が「エージェント  $i$  は、 $\phi$  が真であると信じている」を表す。

- $\phi$  が論理式ならば、 $G\phi$  は論理式とする。また省略形として、以下を定義する。

$$F\phi := \neg G\neg\phi$$

直感的には、 $G\phi$  が「 $\phi$  はこれからの未来、常に真になる」、 $F\phi$  が「 $\phi$  は現在を含む未来のある時点で真になる」を表す。

- $\phi$  が論理式ならば、 $H\phi$  も論理式とする。また省略形として、以下を定義する。

$$P\phi := \neg H\neg\phi$$

直感的には、 $H\phi$  が「 $\phi$  は現在まで常に真であった」、 $P\phi$  が「 $\phi$  は現在を含む過去のある時点で真であった」を表す。

- $\phi$  が論理式、 $i \in AGT$ 、 $e$  が  $0 \leq e \leq 1$  を満たす実数のとき、 $Effort_e^i \phi$  は論理式であるとする。  
直感的には、「エージェント  $i$  は  $\phi$  を真にするため、程度  $e$  の努力をした」を表す。
- $\phi$  が論理式、 $i \in AGT$ 、 $v$  (value) が  $0 \leq v \leq 1$  を満たすとき、 $Deserve_v^i \phi$  は論理式であるとする。  
直感的には、「エージェント  $i$  は程度  $v$  で  $\phi$  に相応しい」を表す。
- $\phi$  が論理式、 $i \in AGT$ 、 $p$  (praiseworthiness) が  $0 \leq p \leq 1$  を満たすとき、 $Praise_p^i \phi$  は論理式であるとする。  
直感的な意味は、「エージェント  $i$  が  $\phi$  を真にすることは、程度  $p$  で評価される」を表す。

この他、オペレータ間に、一般的な優先順位 (例えば、 $\wedge$  が  $\vee$  より先に結合する、 $\rightarrow$  は右結合である、括弧で結合順序を変更できる等) を導入する。

## 5 意味論

### 5.1 構造

以下のものを定めておく。

- 可能世界の集合  $W (\neq \emptyset)$
- 各世界の命題記号に対する真偽の割り当て

$$V : W \times S \rightarrow \mathbb{B}$$

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

$$\mathcal{B}^i : W \times W \rightarrow [0, 1]$$

ただし任意の  $i \in AGT$ ,  $w \in W$  に対し、 $\sum_{w' \in W} \mathcal{B}^i(w, w') = 1$  であること。  
また、 $\mathcal{B}^i$  は、以下の性質を満たすものとする。

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

これはエージェント  $i$  にとっての度合い付きの信念を表す関数である。

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

$$\mathcal{D}^i : W \rightarrow [0, 1]$$

これはエージェント  $i$  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、 $2^W$  への関数

$$\mathcal{G} : W \rightarrow 2^W$$

この  $\mathcal{G}$  は世界間の到達可能性関係とも見ることができる。ただし到達可能性関係  $\mathcal{G}$ ,  $\mathcal{G}^{-1}$  は以下を満たす必要がある。

- 反射的かつ推移的
- $w_1, w_2 \in \mathcal{G}(w)$  ならば、 $w_1 \in \mathcal{G}(w_2)$  もしくは  $w_2 \in \mathcal{G}(w_1)$

これは可能世界  $w$  の未来の可能世界の集合  $\mathcal{G}(w)$  を表す関数である。

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

$$\mathcal{E}^i : \{(w, w') \mid w' \in \mathcal{G}(w)\} \rightarrow [0, 1]$$

これはエージェント  $i$  の努力の度合いを表す関数である。

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

$$\mathcal{V}^i : W \rightarrow [0, 1]$$

これはエージェント  $i$  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

$$\mathcal{P}^i : W \rightarrow [0, 1]$$

これはエージェント  $i$  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。

## 5.2 解釈

論理式  $\phi$  に対し、その解釈  $\llbracket \phi \rrbracket_{\langle M, w \rangle}$  を以下のように定める。

- $P \in ATM$  ならば、 $\llbracket P \rrbracket_{\langle M, w \rangle} = \top$  iff  $V(w, P) = \top$
- $\llbracket \neg \phi \rrbracket_{\langle M, w \rangle} = \top$  iff not  $\llbracket \phi \rrbracket_{\langle M, w \rangle} = \top$
- $\llbracket \phi \vee \psi \rrbracket_{\langle M, w \rangle} = \top$  iff  $\llbracket \phi \rrbracket_{\langle M, w \rangle} = \top$  or  $\llbracket \psi \rrbracket_{\langle M, w \rangle} = \top$
- $\llbracket Prob_{\ell}^i \phi \rrbracket_{\langle M, w \rangle} = \top$  iff  $\sum_{\substack{w' \in W \text{ かつ} \\ w' \text{ で } \phi \text{ が真}}} \mathcal{B}^i(w, w') = \ell$
- $\llbracket G\phi \rrbracket_{\langle M, w \rangle} = \top$  iff  $w' \in \mathcal{G}(w)$  となる任意の  $w'$  に対して、 $\llbracket \phi \rrbracket_{\langle M, w' \rangle} = \top$
- $\llbracket H\phi \rrbracket_{\langle M, w \rangle} = \top$  iff  $w \in \mathcal{G}(w')$  となる任意の  $w'$  に対して、 $\llbracket \phi \rrbracket_{\langle M, w' \rangle} = \top$
- $\llbracket Effort_e^i \phi \rrbracket_{\langle M, w \rangle} = \top$  iff  $w \in \mathcal{G}(w')$  を満たす  $w'$  が存在して、 $\llbracket \phi \rrbracket_{\langle M, w' \rangle} = \top$  and  $\mathcal{E}^i(w, w') = e$
- $\llbracket Des_d^i \phi \rrbracket_{\langle M, w \rangle} = \top$  iff  $\frac{\sum_{\substack{w' \in W \text{ かつ} \\ w' \text{ で } \phi \text{ が真}}} (\mathcal{D}^i(w') \times \mathcal{B}^i(w, w'))}{\sum_{\substack{w' \in W \text{ かつ} \\ w' \text{ で } \phi \text{ が真}}} \mathcal{B}^i(w, w')} = d$
- $\llbracket Deserve_v^i \phi \rrbracket_{\langle M, w \rangle} = \top$  iff  $\frac{\sum_{\substack{w' \in W \text{ かつ} \\ w' \text{ で } \phi \text{ が真}}} (\mathcal{V}^i(w') \times \mathcal{B}^i(w, w'))}{\sum_{\substack{w' \in W \text{ かつ} \\ w' \text{ で } \phi \text{ が真}}} \mathcal{B}^i(w, w')} = v$
- $\llbracket Praise_p^i \phi \rrbracket_{\langle M, w \rangle} = \top$  iff  $\frac{\sum_{\substack{w' \in W \text{ かつ} \\ w' \text{ で } \phi \text{ が真}}} (\mathcal{P}^i(w') \times \mathcal{B}^i(w, w'))}{\sum_{\substack{w' \in W \text{ かつ} \\ w' \text{ で } \phi \text{ が真}}} \mathcal{B}^i(w, w')} = p$

図3の各可視関係に付いている値が  $\mathcal{B}^i(w, w')$  である。この場合、世界  $w'_1, w'_2, w'_3$  で、 $\phi$  が成り立っており、 $\mathcal{B}^i(w, w')$  の合計は、 $0.2 + 0.2 + 0.2 = 0.6$  となり、 $\llbracket Prob_{0.6}^i \phi \rrbracket_{\langle M, w \rangle} = \top$  が成り立つ。また、図4が  $\llbracket Bel^i \phi \rrbracket_{\langle M, w \rangle} = \top$  が成り立つ場合の例である。世界  $w$  から到達可能な世界全てで  $\phi$  が成り立っており、 $\mathcal{B}^i(w, w')$  の合計が1.0である。

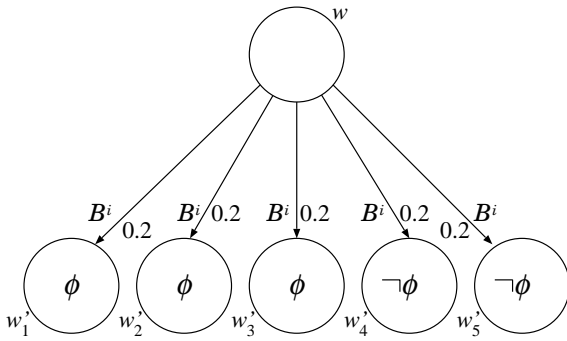


図3: Prob オペレータの解釈

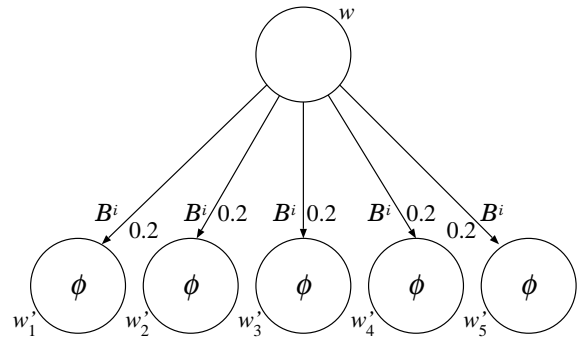


図4: Bel オペレータの解釈

図5の各可視関係に付いている値が  $\mathcal{B}^i(w, w')$ 、各世界に付いている値が  $\mathcal{D}^i(w')$  である。図5では、 $(0.5 \times 0.2 + 0.9 \times 0.2 + 0.1 \times 0.2) / (0.2 + 0.2 + 0.2) = 0.5$  となるので、 $\llbracket Des_{0.5}^i \phi \rrbracket_{\langle M, w \rangle} = \top$  となる。この場合、 $\phi$  の望ましさ 0.5 は、 $\phi$  の成り立つ世界3つの望ましさの平均になっている。同様に、 $\llbracket Des_{0.7}^i(\phi \wedge \psi) \rrbracket_{\langle M, w \rangle} = \top$ 、 $\llbracket Des_{0.3}^i(\phi \wedge \sigma) \rrbracket_{\langle M, w \rangle} = \top$  となる。

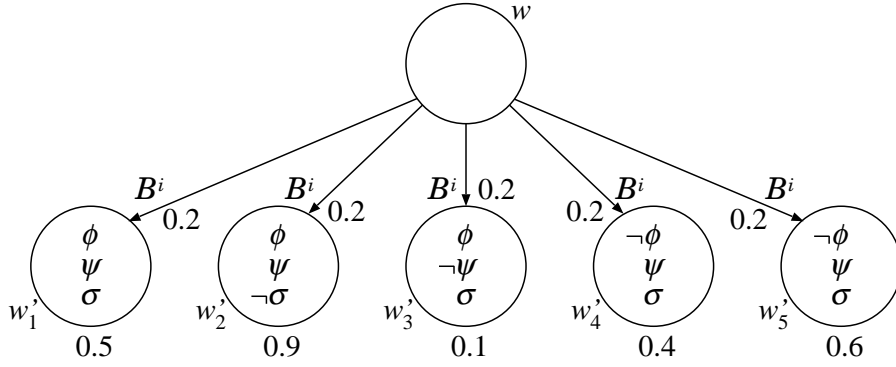


図 5: Des オペレータの解釈

仮に、各可視関係に付いている  $\mathcal{B}^i(w, w')$  の値が、世界  $w'_1$  から順に、0.2, 0.4, 0.2, 0.1, 0.1 の場合、 $\phi$  の望ましさは、 $(0.5 \times 0.2 + 0.9 \times 0.4 + 0.1 \times 0.2) / (0.2 + 0.4 + 0.2) = 0.6$  となり、 $w'_2$  の望ましさに引きずられて高くなる。このように、 $\phi$  の望ましさは、 $\phi$  の成り立つ世界 3 つの望ましさの加重平均になる。

また、この意味論は次のような性質を満たす。[[ $Des_{d_1}^i \phi$ ]] $_{\langle M, w \rangle} = \top$  かつ [[ $Des_{d_2}^i (\phi \wedge \psi)$ ]] $_{\langle M, w \rangle} = \top$  のとき、 $d_2$  は  $d_1$  より大きいことも小さいこともある。すなわち、 $\phi$  より  $\phi \wedge \psi$  の方が望ましいことも、望ましくないこともある。例えば、「単位がとれる」よりも、「“単位がとれる” かつ “卒業できる”」は望ましいが、「“単位がとれる” かつ “留年する”」は望ましくない。

## 6 恒真論理式

我々の論理体系は、本章に示す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なかで使用される。以下において、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。また、 $\frac{\phi}{\psi}$  は、「 $\phi$  が恒真ならば  $\psi$  も恒真である」という性質を表す。

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

$$\Box(\phi \rightarrow \psi) \rightarrow (\Box\phi \rightarrow \Box\psi) \quad (\text{K-}\Box)$$

$$\frac{\phi}{\Box\phi} \quad (\text{RN-}\Box)$$

$$\Box(\phi \wedge \psi) \leftrightarrow (\Box\phi \wedge \Box\psi) \quad (\text{C-}\Box)$$

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、以下が成り立つ。

$$\frac{\phi}{\Diamond\phi} \quad (\text{RN-}\Diamond)$$

$$\Diamond(\phi \wedge \psi) \rightarrow (\Diamond\phi \wedge \Diamond\psi) \quad (\text{C-}\Diamond)$$

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。

$$Bel^i \phi \rightarrow \neg Bel^i \neg \phi \quad (D-Bel^i)$$

$$Bel^i \phi \rightarrow Bel^i Bel^i \phi \quad (4-Bel^i)$$

$$\neg Bel^i \phi \rightarrow Bel^i \neg Bel^i \phi \quad (5-Bel^i)$$

以上の恒真論理式は、エージェントの信念に矛盾がない (D- $Bel^i$ )、エージェントが自身の信念を自覚している (4- $Bel^i$ )、エージェントが自身の信念でないものを信念でないと自覚している (5- $Bel^i$ ) ことを示している。

□ が、オペレータ  $Prob_\ell^i, Des_d^i, Praise_p^i, Deserve_v^i$  のいずれかの場合、以下が成り立つ。これは、エージェントが自身の  $Prob_\ell^i$  や  $Des_d^i$  などの有無に関して自覚していることを示している。

$$\Box \phi \leftrightarrow Bel^i \Box \phi \quad (4-MIX)$$

$$\neg \Box \phi \leftrightarrow Bel^i \neg \Box \phi \quad (5-MIX)$$

## 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、4章で定義したオペレータを用いて、度合い付きの感情を論理式として示す。OCC Theory で定義されている22種類の感情タイプは、2章で記述した通りである。同じグループに属する感情は、生起条件が同じ形で定義されており、「感情の強さに影響する変数」についても、同じ変数が用いられているということに留意されたい。

また、以下で用いる、パラメータを決定する関数 ( $Joy$  の  $f_J$  や、 $Distress$  の  $f_D$  など) は、全て、任意の引数に対し、0以上1以下の値をとるものとする。

### 7.1 Reaction to events クラス

#### Well-being Emotions グループ

##### 7.1.1 Joy

OCC Theory において、感情タイプ  $Joy$  の生起条件は、「pleased about a desirable event」と定義されており、望ましいイベントに対する喜びの感情タイプである。また、強さに影響する変数には、以下の変数が定義されている。

- イベントがどの程度望ましいか

この変数を、望ましさの程度を変数に持つ、オペレータ  $Des$  で表現し、 $Joy$  オペレータを以下のように定義する。

$$Joy_{f_J(d)}^i \phi := Des_d^i \phi \wedge Bel^i \phi$$

関数  $f_J$  は、 $d$  に関する増加関数とする。

### 7.1.2 Distress

OCC Theory において、感情タイプ Distress の生起条件は、「displeased about a undesirable event」と定義されており、望ましくないイベントに対する悲しみの感情タイプである。また、強さに影響する変数には、以下の変数が定義されている。

- イベントがどの程度望ましくないか

この変数を、望ましさの程度を変数に持つ、オペレータ  $Des$  で表現し、 $Distress$  オペレータを以下のように定義する。

$$Distress_{f_D(d)}^i \phi := Des_d^i \phi \wedge Bel^i \phi$$

関数  $f_D$  は、 $d$  に関する減少関数とする。

## Fortunes-of-others Emotions グループ

### 7.1.3 HappyFor

OCC Theory において、感情タイプ HappyFor の生起条件は、「pleased about an event presumed to be desirable for someone else」と定義されており、他者にとって望ましいと思われるイベントに対する喜びの感情タイプである。また、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自分にとってどの程度望ましいか
- イベントが他者にとってどの程度望ましいと思われるか
- 他者がどの程度イベントに相応しいか
- 他者をどの程度好ましく思っているか

変数「他者にとって望ましいイベントが自分にとってどの程度望ましいか」は、論理式  $Des_{d_i}^i Bel^j \phi$  で表現できる。変数「イベントが他者にとってどの程度望ましいと思われるか」は、他者がどの程度イベントを望ましく思うかという、自分自身の推測であるが、「イベントが他者にとってどの程度望ましいか」と考え、見込みを表すオペレータ  $Prob$  は使わず、論理式  $Bel^i Des_{d_j}^j \phi$  で表現する。変数「他者がどの程度イベントに相応しいか」は、イベントに対する相応しさを表すオペレータ  $Deserve$  を用いて、 $Bel^i Deserve_v^j \phi$  のように表現できる。なお、本研究では、エージェント間の友好度を表すオペレータについては、簡単化のため、導入を省いた。この理由は、7.3 節と同様、友好の対象を引数としてとるためには、述語論理が必要となるためである。

$HappyFor$  オペレータを以下のように定義する。

$$HappyFor_{f_{HF}(d_i, d_j, v)}^{i, j} \phi := Bel^i \phi \wedge Des_{d_i}^i Bel^j \phi \wedge Bel^i Des_{d_j}^j \phi \wedge Bel^i Deserve_v^j \phi$$

ここで、関数  $f_{HF}$  は、 $d_i$  に関する増加関数、 $d_j$  に関する増加関数、 $v$  に関する増加関数とする。

### 7.1.4 SorryFor

OCC Theory において、感情タイプ SorryFor の生起条件は、「displeased about an event presumed to be undesirable for someone else」と定義されており、他者にとって望ましくないと思われるイベントに対する悲しみの感情タイプである。また、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自分にとってどの程度望ましくないか

- イベントが他者にとってどの程度望ましくないと思われるか
- 他者がどの程度イベントに相応しくないか
- 他者をどの程度好ましく思っているか

*HappyFor* オペレータと同様、*SorryFor* オペレータを以下のように定義する。

$$SorryFor_{f_{SF}(d_i, d_j, v)}^{i, j} \phi := Bel^i \phi \wedge Des_{d_i}^i Bel^j \phi \wedge Bel^i Des_{d_j}^j \phi \wedge Bel^i Deserve_v^j \phi$$

ここで、関数  $f_{SF}$  は、 $d_i$  に関する減少関数、 $d_j$  に関する減少関数、 $v$  に関する減少関数とする。

### 7.1.5 Resentment

OCC Theory において、感情タイプ Resentment の生起条件は、「displeased about an event presumed to be desirable for someone else」と定義されており、他者にとって望ましいと思われるイベントに対する悲しみの感情タイプである。また、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自分にとってどの程度望ましくないか
- イベントが他者にとってどの程度望ましいと思われるか
- 他者がどの程度イベントに相応しくないか
- 他者をどの程度好ましく思っていないか

*HappyFor* オペレータと同様、*Resentment* オペレータを以下のように定義する。

$$Resentment_{f_{Rs}(d_i, d_j, v)}^{i, j} \phi := Bel^i \phi \wedge Des_{d_i}^i Bel^j \phi \wedge Bel^i Des_{d_j}^j \phi \wedge Bel^i Deserve_v^j \phi$$

ここで、関数  $f_{Rs}$  は、 $d_i$  に関する減少関数、 $d_j$  に関する増加関数、 $v$  に関する減少関数とする。

### 7.1.6 Gloating

OCC Theory において、感情タイプ Gloating の生起条件は、「pleased about an event presumed to be undesirable for someone else」と定義されており、他者にとって望ましくないと思われるイベントに対する喜びの感情タイプである。また、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自分にとってどの程度望ましいか
- イベントが他者にとってどの程度望ましくないと思われるか
- 他者がどの程度イベントに相応しいか
- 他者をどの程度好ましく思っていないか

*HappyFor* オペレータと同様、*Gloating* オペレータを以下のように定義する。

$$Gloating_{f_{Gl}(d_i, d_j, v)}^{i, j} \phi := Bel^i \phi \wedge Des_{d_i}^i Bel^j \phi \wedge Bel^i Des_{d_j}^j \phi \wedge Bel^i Deserve_v^j \phi$$

ここで、関数  $f_{Gl}$  は、 $d_i$  に関する増加関数、 $d_j$  に関する減少関数、 $v$  に関する増加関数とする。

## Prospect-based Emotions グループ

### 7.1.7 Hope

OCC Theory において、感情タイプ Hope の生起条件は、「pleased about the prospect of a desirable event」と定義されており、望ましいイベントの可能性に対する喜びの感情タイプである。また、強さに影響する変数には、以下の変数が定義されている。

- イベントがどの程度望ましいか
- イベントが起こる可能性

変数「イベントがどの程度望ましいか」は、論理式  $Des_d^i\phi$  で表現し、変数「イベントが起こる可能性」は、イベントに対するエージェントの見込みを表すオペレータを用いて、論理式  $Prob_\ell^i\phi$  で表現する。ここで、イベントに対する見込みの度合いを表すオペレータの変数  $\ell$  は、1.0 でないとする。 $Prob_{1.0}^i\phi$  の省略形を  $Bel^i\phi$  と定義したので、 $\ell$  が 1.0 だと、現在の信念となり、見込みではなくなる。

*Hope* オペレータを以下のように定義する。

$$Hope_{f_H(d,\ell)}^i\phi := Des_d^i\phi \wedge Prob_\ell^i\phi \wedge \ell \neq 1$$

関数  $f_H$  は、 $d$  に関する増加関数、 $\ell$  に関する増加関数とする。。

### 7.1.8 Fear

OCC Theory において、感情タイプ Fear の生起条件は、「displeased about the prospect of an undesirable event」と定義されており、望ましくないイベントの可能性に対する悲しみの感情タイプである。また、強さに影響する変数には、以下の変数が定義されている。

- イベントがどの程度望ましくないか
- イベントが起こる可能性

*Hope* と同様、*Fear* オペレータを以下のように定義する。

$$Fear_{f_F(d,\ell)}^i\phi := Des_d^i\phi \wedge Prob_\ell^i\phi \wedge \ell \neq 1$$

関数  $f_F$  は、 $d$  に関する減少関数、 $\ell$  に関する増加関数とする。

### 7.1.9 Satisfaction

OCC Theory において、感情タイプ Satisfaction の生起条件は、「pleased about the confirmation of the prospect of a desirable event」と定義されており、望ましいイベントの可能性を確認したことに対する喜びの感情タイプである。また、強さに影響する変数には、以下の変数が定義されている。

- 付随する Hope の強さ
- イベントを達成するために費やされた努力の程度
- イベントの実現度

変数「付随する Hope の強さ」は、論理式  $Bel^i PHope_x^i \phi$  で表現し、変数「イベントを達成するために費やされた努力の程度」は、 $Bel^i PEffort_e^i \phi$  で表現できる。変数「イベントの実現度」は、イベントに対する見込みと考え、 $Prob_\ell^i \phi$  と表現する。

*Satisfaction* オペレータを以下のように定義する。

$$Satisfaction_{f_{St}(x,e,\ell)}^i \phi := Bel^i PHope_x^i \phi \wedge Bel^i PEffort_e^i \phi \wedge Prob_\ell^i \phi$$

関数  $f_{St}$  は、 $x$  に関する増加関数、 $e$  に関する増加関数、 $\ell$  に関する増加関数とする。

### 7.1.10 FearConfirmed

OCC Theory において、感情タイプ FearConfirmed の生起条件は、「displeased about the confirmation of the prospect of an undesirable event」と定義されており、望ましくないイベントの可能性を確認したことに対する悲しみの感情タイプである。また、強さに影響する変数には、以下の変数が定義されている。

- 付随する Fear の強さ
- イベントを回避するために費やされた努力の程度
- イベントの実現度

*Satisfaction* オペレータと同様に、*FearConfirmed* オペレータを以下のように定義する。

$$FearConfirmed_{f_{FC}(x,e,\ell)}^i \phi := Bel^i PFear_x^i \phi \wedge Bel^i PEffort_e^i \neg \phi \wedge Prob_\ell^i \phi$$

関数  $f_{FC}$  は、 $x$  に関する増加関数、 $e$  に関する増加関数、 $\ell$  に関する増加関数とする。

### 7.1.11 Relief

OCC Theory において、感情タイプ Relief の生起条件は、「pleased about the disconfirmation of the prospect of an undesirable event」と定義されており、望ましくないイベントの可能性を確認しなかったことに対する喜びの感情タイプである。また、強さに影響する変数には、以下の変数が定義されている。

- 付随する Fear の強さ
- イベントを回避するために費やされた努力の程度
- イベントの実現度

*Satisfaction* オペレータと同様に、*Relief* オペレータを以下のように定義する。

$$Relief_{f_{RI}(x,e,\ell)}^i \phi := Bel^i PFear_x^i \neg \phi \wedge Bel^i PEffort_e^i \phi \wedge Prob_\ell^i \phi$$

関数  $f_{RI}$  は、 $x$  に関する増加関数、 $e$  に関する増加関数、 $\ell$  に関する増加関数とする。

### 7.1.12 Disappointment

OCC Theory において、感情タイプ Disappointment の生起条件は、「displeased about the disconfirmation of the prospect of an desirable event」と定義されており、望ましいイベントの可能性を確認しなかったことに対する悲しみの感情タイプである。また、強さに影響する変数には、以下の変数が定義されている。

- 付随する Hope の強さ
- イベントを達成するために費やされた努力の程度
- イベントの実現度

*Satisfaction* オペレータと同様に、*Disappointment* オペレータを以下のように定義する。

$$Disappointment_{f_{Ds}(x,e,\ell)}^i \phi := Bel^i P Hope_x^i \neg \phi \wedge Bel^i P Effort_e^i \neg \phi \wedge Prob_\ell^i \phi$$

関数  $f_{Ds}$  は、 $x$  に関する増加関数、 $e$  に関する増加関数、 $\ell$  に関する増加関数とする。

## 7.2 Reaction to agents クラス

### Attribution Emotions グループ

#### 7.2.1 Pride

OCC Theory において、感情タイプ Pride の生起条件は、「approving of one's own praiseworthy action」と定義されており、自分自身の行動に対する自己の称賛の評価である。また、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。

- どの程度称賛されるか
- 認知単位の強さ
- エージェントの行動の意外性

変数「イベントがどの程度称賛されるか」は、イベントの評価を表すオペレータを用いて、論理式  $Praise_p^i \phi$  と記述する。変数「エージェントの行動の意外性」は、 $Bel^i P Prob_\ell^i \phi \wedge Bel^i \phi$  と表現する。この論理式は、直感的には、「過去に見込みが低いと思われていたことが現在、真である」という意味になる。変数「認知単位の強さ」については、エージェントが他エージェントを同一視する程度を表すもので、エージェント間の関係を表すオペレータについては、7.1.3 と同様の理由で省略した。

*Pride* オペレータを以下のように定義する。

$$Pride_{f_{Pr}(p,\ell)}^i \phi := Praise_p^i \phi \wedge Bel^i P Prob_\ell^i \phi \wedge Bel^i \phi$$

関数  $f_{Pr}$  は、 $\ell$  に関する減少関数、 $p$  に関する増加関数とする。

#### 7.2.2 Shame

OCC Theory において、感情タイプ Shame の生起条件は、「disapproving of one's own blame-worthy action」と定義されており、自分自身の行動に対する自己の非難の評価である。また、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。

- どの程度非難されるか
- 認知単位の強さ
- エージェントの行動の意外性

*Pride* オペレータと同様に、*Shame* オペレータを以下のように定義する。

$$Shame_{f_{Sh}(p,\ell)}^i \phi := Praise_p^i \phi \wedge Bel^i P Prob_\ell^i \phi \wedge Bel^i \phi$$

関数  $f_{Sh}$  は、 $\ell$  に関する減少関数、 $p$  に関する減少関数とする。

### 7.2.3 Admiration

OCC Theory において、感情タイプ Admiration の生起条件は、「approving of someone else's praiseworthy action」と定義されており、他者の行動に対する称賛の評価である。また、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。

- どの程度称賛されるか
- エージェントの行動の意外性

変数「イベントがどの程度称賛されるか」については、他エージェント  $j$  がイベント  $\phi$  を真にすることに對し、自エージェント  $i$  が評価するので、論理式  $Bel^i Praise_p^j \phi$  で表す。変数「エージェントの行動の意外性」については、論理式  $Bel^i PProb_\ell^i \phi$  で表すことができる。

*Admiration* オペレータを以下のように定義する。

$$Admiration_{f_{Ad}(p,\ell)}^{i,j} \phi := Bel^i Praise_p^j \phi \wedge Bel^i PProb_\ell^i \phi \wedge Bel^i \phi$$

関数  $f_{Ad}$  は、 $\ell$  に関する減少関数、 $p$  に関する増加関数とする。

### 7.2.4 Reproach

OCC Theory において、感情タイプ Reproach の生起条件は、「approving of someone else's blameworthy action」と定義されており、他者の行動に対する非難の評価である。また、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。

- どの程度非難されるか
- エージェントの行動の意外性

*Admiration* オペレータと同様に、*Reproach* オペレータを以下のように定義する。

$$Reproach_{f_{Rp}(p,\ell)}^{i,j} \phi := Bel^i Praise_p^j \phi \wedge Bel^i PProb_\ell^i \phi \wedge Bel^i \phi$$

関数  $f_{Rp}$  は、 $\ell$  に関する減少関数、 $p$  に関する減少関数とする。

## Well-being/Attribution compounds Emotion グループ

このグループに属する感情タイプは、Well-being Emotions と Attribution Emotions の混合型である。生起条件、変数ともに、Well-being Emotions と Attribution Emotions で定義されたもので構成されており、既に定義した感情のオペレータを用いて、それぞれ新たに感情が定義できる。

### 7.2.5 Gratification

OCC Theory において、感情タイプ Gratification の生起条件は、「approving of one's own praiseworthy action and being pleased about the related desirable event」と定義されており、自分の行動に対する称賛の評価と、それに関連する望ましいイベントに対する喜びである。また、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。

- どの程度称賛されるか
- 認知単位の強さ

- エージェントの行動の意外性
- イベントがどの程度望ましいか

「どの程度称賛されるか」、「認知単位の強さ」、「エージェントの行動の意外性」の3つの変数は、感情タイプ *Pride* で定義されている変数と同じであり、「イベントがどの程度望ましいか」という変数は、感情タイプ *Joy* の変数と同じである。

*Gratification* オペレータは、*Joy* オペレータと *Pride* オペレータを用いて、以下のように記述することができる。

$$Gratification_{f_{Grtf}(x,y)}^i \phi := Joy_x^i \phi \wedge Pride_y^i \phi$$

関数  $f_{Grtf}$  は、 $x$  に関する増加関数、 $y$  に関する増加関数とする。

### 7.2.6 Remorse

OCC Theory において、感情タイプ *Remorse* の生起条件は、「disapproving of one's own blame-worthy action and being displeased about the related undesirable event」と定義されており、自分の行動に対する非難の評価と、それに関連する望ましくないイベントに対する悲しみの感情である。また、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。

- どの程度非難されるか
- 認知単位の強さ
- エージェントの行動の意外性
- イベントがどの程度望ましくないか

「どの程度非難されるか」、「認知単位の強さ」、「エージェントの行動の意外性」の3つの変数は、感情タイプ *Shame* で定義されている変数と同じであり、「イベントがどの程度望ましくないか」という変数は、感情タイプ *Distress* の変数と同じである。

*Remorse* オペレータは、*Distress* オペレータと *Shame* オペレータを用いて、以下のように記述することができる。

$$Remorse_{f_{Rm}(x,y)}^i \phi := Distress_x^i \phi \wedge Shame_y^i \phi$$

関数  $f_{Rm}$  は、 $x$  に関する増加関数、 $y$  に関する増加関数とする。

### 7.2.7 Gratitude

OCC Theory において、感情タイプ *Gratitude* の生起条件は、「approving of someone else's praiseworthy action and being pleased about the related desirable event」と定義されており、他者の行動に対する称賛の評価と、それに関連する望ましいイベントに対する喜びである。また、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。

- どの程度称賛されるか
- エージェントの行動の意外性
- イベントがどの程度望ましいか

「どの程度称賛されるか」、「エージェントの行動の意外性」の2つの変数は、感情タイプ *Admiration* で定義されている変数と同じであり、「イベントがどの程度望ましいか」という変数は、感情タイプ *Joy* の変数と同じである。

*Gratitude* オペレータは、*Joy* オペレータと *Admiration* オペレータを用いて、以下のように記述することができる。

$$Gratitude_{f_{Grtt}(x,y)}^{i,j} \phi := Joy_x^i \phi \wedge Admiration_y^{i,j} \phi$$

関数  $f_{Grtt}$  は、 $x$  に関する増加関数、 $y$  に関する増加関数とする。

### 7.2.8 Anger

OCC Theory において、感情タイプ *Anger* の生起条件は、「disapproving of someone else's blameworthy action and being displeased about the related undesirable event」と定義されており、他者の行動に対する非難の評価と、それに関連する望ましくないイベントに対する悲しみである。また、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。

- どの程度非難されるか
- エージェントの行動の意外性
- イベントがどの程度望ましくないか

「どの程度非難されるか」、「エージェントの行動の意外性」の2つの変数は、感情タイプ *Reproach* で定義されている変数と同じであり、「イベントがどの程度望ましくないか」という変数は、感情タイプ *Distress* の変数と同じである。*Anger* オペレータは、*Distress* オペレータと *Reproach* オペレータを用いて、以下のように記述することができる。

$$Anger_{f_{An}(x,y)}^{i,j} \phi := Distress_x^i \phi \wedge Reproach_y^{i,j} \phi$$

関数  $f_{An}$  は、 $x$  に関する増加関数、 $y$  に関する増加関数とする。

## 7.3 Reaction to objects クラス

### Attraction emotions

このグループに属する感情タイプは、Adam らの形式化からは除かれている。その理由として、このグループに属する2個の感情タイプを形式化するには、引数として対象物をとるため、一階述語論理が必要になると、Adam らは述べている。今回の我々の論理体系も述語論理は扱えないため、このグループの感情タイプについては、今回、考慮から外すこととする。

## 8 感情に関する性質の形式化

本章では、本論文で提案した論理体系の妥当性を示すため、感情に関して成り立つ、もっともらしい性質を表す論理式を示す。ここで成り立つ恒真論理式の証明は、容易なものについては省略し、8.2.1 と 8.4.1 については、付録に証明を記載する。

なお、 $\vdash \phi$  は、「 $\phi$  が恒真である」を表す。

## 8.1 感情同士の対立

### 8.1.1 対立関係の感情

同じグループに属する、ポジティブな感情とネガティブな感情、つまり、相反するような感情は、同時に生じられないという性質である。

- *Joy* の  $f_J$  が、 $f_J(d)=1.0$  iff  $d=1.0$  を満たし、かつ、  
*Distress* の  $f_D$  が、 $f_D(d)=1.0$  iff  $d=0$  を満たすとき、  
 $\vdash \neg(\text{Joy}_{1.0}^i\phi \wedge \text{Distress}_{1.0}^i\phi)$
- *HappyFor* の  $f_{HF}$  と、*SorryFor* の  $f_{SF}$  が、以下のいずれかを満たすとき、
  - $f_{HF}(d_i, d_j, v)=1.0$  iff  $d_i=1.0$  かつ、 $f_{SF}(d_i, d_j, v)=1.0$  iff  $d_i=0$
  - $f_{HF}(d_i, d_j, v)=1.0$  iff  $d_j=1.0$  かつ、 $f_{SF}(d_i, d_j, v)=1.0$  iff  $d_j=0$
  - $f_{HF}(d_i, d_j, v)=1.0$  iff  $v=1.0$  かつ、 $f_{SF}(d_i, d_j, v)=1.0$  iff  $v=0$ $\vdash \neg(\text{HappyFor}_{1.0}^{i,j}\phi \wedge \text{SorryFor}_{1.0}^{i,j}\phi)$
- *Resentment* の  $f_{Rs}$  と、*Gloating* の  $f_{Gl}$  が、以下のいずれかを満たすとき、
  - $f_{Rs}(d_i, d_j, v)=1.0$  iff  $d_i=0$  かつ、 $f_{Gl}(d_i, d_j, v)=1.0$  iff  $d_i=1.0$
  - $f_{Rs}(d_i, d_j, v)=1.0$  iff  $d_j=1.0$  かつ、 $f_{Gl}(d_i, d_j, v)=1.0$  iff  $d_j=0$
  - $f_{Rs}(d_i, d_j, v)=1.0$  iff  $v=0$  かつ、 $f_{Gl}(d_i, d_j, v)=1.0$  iff  $v=1.0$ $\vdash \neg(\text{Resentment}_{1.0}^{i,j}\phi \wedge \text{Gloating}_{1.0}^{i,j}\phi)$
- *Hope* の  $f_H$  が、 $f_H(d)=1.0$  iff  $d=1.0$  を満たし、かつ、  
*Fear* の  $f_F$  が、 $f_F(d)=1.0$  iff  $d=0$  を満たすとき、  
 $\vdash \neg(\text{Hope}_{1.0}^i\phi \wedge \text{Fear}_{1.0}^i\phi)$
- *Satisfaction* の  $f_{St}$  が、 $f_{St}(x, e, \ell)=1.0$  iff  $x=1.0$  を満たし、かつ、  
*FearConfirmed* の  $f_{FC}$  が、 $f_{FC}(x, e, \ell)=1.0$  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、  
 $\vdash \neg(\text{Satisfaction}_{1.0}^i\phi \wedge \text{FearConfirmed}_{1.0}^i\phi)$
- *Relief* の  $f_{Rl}$  が、 $f_{Rl}(x, e, \ell)=1.0$  iff  $x=1.0$  を満たし、かつ、  
*Disappointment* の  $f_{Ds}$  が、 $f_{Ds}(x, e, \ell)=1.0$  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、  
 $\vdash \neg(\text{Relief}_{1.0}^i\phi \wedge \text{Disappointment}_{1.0}^i\phi)$
- *Pride* の  $f_{Pr}$  が、 $f_{Pr}(p, \ell)=1.0$  iff  $p=1.0$  を満たし、かつ、  
*Shame* の  $f_{Sh}$  が、 $f_{Sh}(p, \ell)=1.0$  iff  $p=0$  を満たすとき、  
 $\vdash \neg(\text{Pride}_{1.0}^i\phi \wedge \text{Shame}_{1.0}^i\phi)$
- *Admiration* の  $f_{Ad}$  が、 $f_{Ad}(p, \ell)=1.0$  iff  $p=1.0$  を満たし、かつ、  
*Reproach* の  $f_{Rp}$  が、 $f_{Rp}(p, \ell)=1.0$  iff  $p=0$  を満たすとき、  
 $\vdash \neg(\text{Admiration}_{1.0}^{i,j}\phi \wedge \text{Reproach}_{1.0}^{i,j}\phi)$
- *Gratification* の  $f_{Grtf}$  と、*Remorse* の  $f_{Rm}$  が、以下のいずれかを満たすとき、
  - $f_{Grtf}(x, y)=1.0$  iff  $x=1.0$  かつ、 $f_{Rm}(x, y)=1.0$  iff  $x=1.0$  かつ、  
上記の *Joy* と *Distress* に関する条件が満たされる

- $f_{Grtf}(x, y)=1.0$  iff  $y=1.0$  かつ、 $f_{Rm}(x, y)=1.0$  iff  $y=1.0$  かつ、  
上記の *Pride* と *Shame* に関する条件が満たされる

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

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

$$\vdash \neg(\text{Gratitude}_{1.0}^{i,j} \phi \wedge \text{Anger}_{1.0}^{i,j} \phi)$$

### 8.1.2 友好感情と嫌悪感情

友好的な感情と、嫌悪的な感情は同時に生起されないという性質である。エージェントは、あるエージェントと、ある事象に対し、友好的な感情と嫌悪的な感情を同時に持たないということを証明できる。

- *HappyFor* の  $f_{HF}$  と、*Resentment* の  $f_{Rs}$  が、以下のいずれかを満たすとき、
  - $f_{HF}(d_i, d_j, v)=1.0$  iff  $d_i=1.0$  かつ、 $f_{Rs}(d_i, d_j, v)=1.0$  iff  $d_i=0$
  - $f_{HF}(d_i, d_j, v)=1.0$  iff  $v=1.0$  かつ、 $f_{Rs}(d_i, d_j, v)=1.0$  iff  $v=0$

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

- *HappyFor* の  $f_{HF}$  と、*Gloating* の  $f_{Gl}$  が、以下を満たすとき、
  - $f_{HF}(d_i, d_j, v)=1.0$  iff  $d_j=1.0$  かつ、 $f_{Gl}(d_i, d_j, v)=1.0$  iff  $d_j=0$

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

- *SorryFor* の  $f_{SF}$  と、*Resentment* の  $f_{Rs}$  が、以下を満たすとき、
  - $f_{SF}(d_i, d_j, v)=1.0$  iff  $d_j=0$  かつ、 $f_{Rs}(d_i, d_j, v)=1.0$  iff  $d_j=1.0$

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

- *SorryFor* の  $f_{SF}$  と、*Gloating* の  $f_{Gl}$  が、以下のいずれかを満たすとき、
  - $f_{SF}(d_i, d_j, v)=1.0$  iff  $d_i=0$  かつ、 $f_{Gl}(d_i, d_j, v)=1.0$  iff  $d_i=1.0$
  - $f_{SF}(d_i, d_j, v)=1.0$  iff  $v=0$  かつ、 $f_{Gl}(d_i, d_j, v)=1.0$  iff  $v=1.0$

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

## 8.2 Prospect-based Emotions とその確証性

### 8.2.1 Confirmation Emotions と Well-being Emotions の関係

Confirmation Emotions が事象  $\phi$  に対して生起されたとき、その事象の望ましさが過去と現在で変化がない場合、その望ましさに伴う Well-being Emotions が生起される。ポジティブな Confirmation Emotions は Joy を誘発し、ネガティブな Confirmation Emotions は Distress を誘発する、という性質である。

- *Satisfaction* の  $f_{St}$  が、 $f_{St}(x, e, \ell)=1.0$  iff  $\ell=1.0$  を満たし、かつ、任意の  $d$  について、 $Bel^i(PDes_d^i\phi \rightarrow Des_d^i\phi)$  が真のとき、ある  $d$  が存在して、  
 $\vdash Satisfaction_{1.0}^i\phi \rightarrow Joy_{f_J(d)}^i\phi$
- *FearConfirmed* の  $f_{FC}$  が、 $f_{FC}(x, e, \ell)=1.0$  iff  $\ell=1.0$  を満たし、かつ、任意の  $d$  について、 $Bel^i(PDes_d^i\phi \rightarrow Des_d^i\phi)$  が真のとき、ある  $d$  が存在して、  
 $\vdash FearConfirmed_{1.0}^i\phi \rightarrow Distress_{f_D(d)}^i\phi$
- *Relief* の  $f_{RI}$  が、 $f_{RI}(x, e, \ell)=1.0$  iff  $\ell=1.0$  を満たし、かつ、任意の  $d$  について、 $Bel^i(PDes_d^i\phi \rightarrow Des_d^i\phi)$  が真のとき、ある  $d$  が存在して、  
 $\vdash Relief_{1.0}^i\phi \rightarrow Joy_{f_J(d)}^i\phi$
- *Disappointment* の  $f_{Ds}$  が、 $f_{Ds}(x, e, \ell)=1.0$  iff  $\ell=1.0$  を満たし、かつ、任意の  $d$  について、 $Bel^i(PDes_d^i\phi \rightarrow Des_d^i\phi)$  が真のとき、ある  $d$  が存在して、  
 $\vdash Disappointment_{1.0}^i\phi \rightarrow Distress_{f_D(d)}^i\phi$

## 8.2.2 予想に対する確認と未確認の同時性

予想した出来事が、現実になったときに生起される感情と、現実にならなかったときに生起される感情は、同時には成り立たないという性質である。

- *Satisfaction* の  $f_{St}$  が、 $f_{St}(x, e, \ell)=1.0$  iff  $\ell=1.0$  を満たし、かつ、  
*Disappointment* の  $f_{Ds}$  が、 $f_{Ds}(x, e, \ell)=1.0$  iff  $\ell=1.0$  を満たすとき、  
 $\vdash \neg(Satisfaction_{1.0}^i\phi \wedge Disappointment_{1.0}^i\neg\phi)$
- *FearConfirmed* の  $f_{FC}$  が、 $f_{FC}(x, e, \ell)=1.0$  iff  $\ell=1.0$  を満たし、かつ、  
*Relief* の  $f_{RI}$  が、 $f_{RI}(x, e, \ell)=1.0$  iff  $\ell=1.0$  を満たすとき、  
 $\vdash \neg(FearConfirmed_{1.0}^i\phi \wedge Relief_{1.0}^i\neg\phi)$

## 8.3 自分自身と他者の Attribution Emotions

### 8.3.1 自分自身に対する他者の感情

他者の行動の評価に関する感情が、自身から自身に向けられたとき、他者の行動の評価に関する感情は、自分自身の感情と同等であるという性質である。

- *Admiration* の  $f_{Ad}$  と、*Pride* の  $f_{Pr}$  が同じ関数であるとき、  
 $\vdash Admiration_{f_{Ad}(p,\ell)}^{i,i}\phi \leftrightarrow Pride_{f_{Pr}(p,\ell)}^i\phi$
- *Reproach* の  $f_{Rp}$  と、*Shame* の  $f_{Sh}$  が同じ関数であるとき、  
 $\vdash Reproach_{f_{Rp}(p,\ell)}^{i,i}\phi \leftrightarrow Shame_{f_{Sh}(p,\ell)}^i\phi$

### 8.3.2 他者の感情が自身に与える影響

他者に自身の行動を評価されたときに、自身がその感情に伴う感情を生起するとは限らないという性質が成り立つ。例えば、誰かが自分に対し *Admiration* を感じて、自分が *Pride* を感じることはないかもしれない。

$$\not\vdash Bel^i Admiration_{f_{Ad}(p,\ell)}^{j,i} \phi \rightarrow Pride_{f_{Pr}(p,\ell)}^i \phi$$

$$\not\vdash Bel^i Reproach_{f_{Rp}(p,\ell)}^{j,i} \phi \rightarrow Shame_{f_{Sh}(p,\ell)}^i \phi$$

## 8.4 その他の性質

### 8.4.1 感情の自覚

定義した 20 個全ての感情 *Emotion* について、以下が成り立つ。

$$\vdash Emotion_f^i \phi \leftrightarrow Bel^i Emotion_f^i \phi$$

$$\vdash \neg Emotion_f^i \phi \leftrightarrow Bel^i \neg Emotion_f^i \phi$$

### 8.4.2 感情の生起に必要な条件

全ての感情は、定義にオペレータ  $Des_d^i$  と  $Praise_p^i$  の少なくとも一方を用いている。つまり、全ての感情は、ある程度の度合いで、願望あるいは称賛を持たなければ生じないということである。例えば、*Joy* の  $f_J$  が、「 $d = 0$  ならば、 $f_J(d) = 0$ 」を満たすならば、いかなる  $\phi$  に対しても  $Des_0^i \phi$  を満たすようなエージェントは、いかなる  $\phi$  に対しても  $Joy_0^i \phi$  を満たす。すなわち、 $Des_d^i$  を持たないエージェントには *Joy* は生じないということである。

これは、Adam らの形式化で述べられている、「願望も理想も持たないエージェントは感情を持たない」という性質の一般化である。

## 9 考察

本章では、今回、我々が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。

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。それに倣い、我々の論理体系の妥当性についても、8 章で、感情に関する性質の形式化によって示した。我々の論理体系は、Adam らの論理体系に準拠し、新たにオペレータと変数を導入し、拡張したものである。Adam らの論理体系で成り立つ、感情に関する自然な性質は、我々の論理体系でもほぼ成り立つと言える。しかし、Adam らの論理体系とは異なり、我々の論理体系は、感情の強度について考慮しなければならず、感情の性質を表現する場合には、変数と関数について制限を設ける必要がある。

8.1.1 節で述べた、対立関係にある感情は、Adam らの論理体系ではどのような場合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、2 つの感情の強度が最大という条件のもとでは同時に生起しないが、相反する感情でも、どちらの強度も最大でないような場合には、同時に生起されることは有り得る。例えば、Adam らの論理体系では  $\vdash \neg (Joy^i \phi \wedge Distress^i \phi)$  が成り立つが、我々の論理体系では感情の強度を扱うため、 $\vdash \neg (Joy_{1.0}^i \phi \wedge Distress_{1.0}^i \phi)$  のような感情の強度が 1.0 となるもの同士では成り立つ。しかし、感情の強度が最大でない場合、例えば、 $Joy_{0.6}^i \phi \wedge Distress_{0.3}^i \phi$  のような論理式は真となることがある。

また、Adam らの論理体系では成り立つ性質が、我々の論理体系では成り立たない場合がある。例えば、 $\neg (Hope^i \phi \wedge Fear^i \neg \phi)$  は、Adam らの論理体系では、恒真であり、これは、 $\phi$  と  $\neg \phi$  のどちらか可能性が高い方に依存して、*Hope* か *Fear* の高々一方のみが生起されることを表す。しかし、この性質は少々、自然でないように思われる。実際、 $\phi$  を期待する一方で、 $\neg \phi$  を不安に思うという状態は、有り得ると考えられる。これに対し、我々の論理体系では、 $Hope_{0.3}^i \phi \wedge Fear_{0.7}^i \neg \phi$

が真になりうる。この方が自然であると考えられる。 $\phi$ に対する望ましさが高くても、現実になる見込みが低ければ、Hopeが生起されたとしても、強度は小さくなるし、 $\neg\phi$ に対する望ましさが0であっても、 $\neg\phi$ が現実になる見込みが少しでもあれば、Fearは生起されるだろう。

以上のように、感情に強度を導入することで、Adamらの論理体系より、柔軟性があり、より人間らしい感情表現が可能になる。

また、我々の形式化は石川[10]によって、Jason[11]で実装された。この実装によって、我々の論理体系が、エージェントの強度付きの感情表現と、生起した感情によって行動を決定できるエージェントの実現が可能であることが示された。

現在の課題は、我々の論理体系で、論理式として表現した感情が、どの程度、実際の人間に近いかという、もっともらしさについて検討していないという点である。この、感情に関する妥当性の検討は、人間の感情という複雑なものを扱う上で、常に考えなければいけない問題である。

また、今回は簡単化のため、OCC Theoryで定義されている感情の強さに影響する評価変数のうち、エージェント間の関係を表す変数 (Fortunes-of-others Emotionsの変数「他者をどの程度好ましく思っているか」、Attribution Emotionsの変数「認知単位の強さ」等) と、Attraction Emotionsで定義されている変数(「対象がどの程度魅力的か」、「対象に対しどの程度親近感を持つか」)を考慮から外した。これらの表現を可能にする方法を、一階述語論理ではなく、新たなオペレータを導入することで、表現可能にする方法の検討も今後の課題である。

## 10 終わりに

本論文では、OCC Theoryで定義されている感情タイプ22種類のうちAttraction Emotionsを除く20種類を、論理式として形式化したAdamらの論理体系をもとに、感情の強さを表現可能にするため、新たなオペレータとそれらオペレータに変数を導入して拡張することで、Adamらの論理体系では表現できなかった、感情の強さを表現できる論理体系を提案した。

Adamらの論理体系で成り立つ、感情に関するもっともらしい性質は、我々の論理体系でも成り立ち、また、感情に強度を導入することで、Adamらの論理体系では表現できなかった、人間らしい、曖昧な表現も可能になった。

今後は、9章で述べた、諸問題に関する検討を進め、よりOCC Theoryに忠実になるよう、徹底した論理体系に近づけることが課題である。

## 謝辞

本論文の執筆及び研究にあたり、指導教官の新出尚之准教授にはいつも丁寧なご指導と助言を賜りました。心から感謝の気持ちとお礼を申し上げます、謝辞にさせていただきます。

## 参考文献

- [1] Andrew Ortony, Gerald L. Clore, and Allan Collins, “The Cognitive Structure of Emotions”, Cambridge University Press, 1988.
- [2] Lazarus, R.S., “Emotion and Adaption”, Oxford University Press, 1991.
- [3] Chellas, B.F., “Modal Logic: an Introduction”, Cambridge University Press, 1980.

- [4] Carole Adam, Andreas Herzig, and Dominique Longin, “A logical formalization of OCC theory of emotions”, *Synthese*, Vol.168, Issue 2, pp.201-248, May 2009.
- [5] Bas R. Steunebrink, Mehdi Dastani, John-Jules Ch. Meyer “A formal model of emotion triggers: an approach for BDI agents”, *Synthese*, Vol.185, Issue 1 Supplement, pp 83-129, April 2012.
- [6] David Pereira, Eugenio Oliveira, and Nelma Moreira, “Modelling emotional BDI agents”, *Workshop on Formal Approaches to Multi-Agent Systems(FAMAS 2006)*, Riva Del Grada, 2006.
- [7] 藤田恵, 清水詩子, 新出尚之, “感情表現を行動決定に取り入れた BDI エージェントの実現に向けて”, *JAWS2011 予稿集*, 2011.
- [8] 池之内彰子, “OCC Theory に基づくエージェントの感情表現の改良”, *奈良女子大学理学部情報科学科 2012 年度卒業論文*, 2013.
- [9] 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”, *奈良女子大学理学部情報科学科 2013 年度卒業論文*, 2014.
- [10] 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”, *奈良女子大学理学部情報科学科 2014 年度卒業論文*, 2015.
- [11] Rafael H. Bordini, Jomi Fred Hübner, and Michael Wooldridge, “Programming Multi-Agent Systems in AgentSpeak using Jason”, *John Wiley & Sons*, 2007.
- [12] 大平英樹, “感情心理学・入門”, *有斐閣*, 2010.

## 付録

ここでは、8章で記述した、感情に関する性質を表す恒真論理式の証明を示す。

### 8.2.1 Confirmation emotions と Well-being emotions の関係

*Satisfaction* の  $f_{St}$  が、 $f_{St}(x, e, \ell)=1.0$  iff  $\ell=1.0$  を満たし、かつ、任意の  $d$  について、 $Bel^i(PDes_d^i\phi \rightarrow Des_d^i\phi)$  が真のとき、ある  $d$  が存在して、 $\vdash Satisfaction_{1.0}^i\phi \rightarrow Joy_{f_J(d)}^i\phi$

#### 証明

*Satisfaction* $_{1.0}^i\phi$  が真かつ、任意の  $d$  について、 $Bel^i(PDes_d^i\phi \rightarrow Des_d^i\phi)$  が真だと仮定して、 $Joy_{f_J(d)}^i\phi$  が真であることを示す。

仮定から、*Satisfaction* $_{1.0}^i\phi$  は真

*Satisfaction* の定義から、*Prob* $_{1.0}^i\phi$  は真、つまり、 $Bel^i\phi$  が真

また、ある  $x$  が存在して、 $Bel^iPHope_x^i\phi$  は真

つまり、ある  $d$  と  $\ell$  が存在して、 $Bel^iP(Des_d^i\phi \wedge Prob_\ell^i\phi)$  が真

$Bel^iP(Des_d^i\phi \wedge Prob_\ell^i\phi)$  が真ならば、

(C-◇)、(RN-□)、(K-□)、(C-□) から、

$Bel^iPDes_d^i\phi \wedge Bel^iPProb_\ell^i\phi$  が真なので、 $Bel^iPDes_d^i\phi$  は真

仮定から、 $Bel^i(PDes_d^i\phi \rightarrow Des_d^i\phi)$  は真

(K-□) から、

$Bel^iPDes_d^i\phi \rightarrow Bel^iDes_d^i\phi$  が真なので、 $Bel^iDes_d^i\phi$  は真

(4-MIX) から、 $Des_d^i\phi$  は真

以上より、 $Bel^i\phi$  と  $Des_d^i\phi$  が真

つまり、 $Bel^i\phi \wedge Des_d^i\phi$  が真

*Joy* の定義より、 $Joy_{f_J(d)}^i\phi$  が真

### 8.4.1 感情の自覚

$$\vdash Emotion_{f_f}^i\phi \leftrightarrow Bel^iEmotion_{f_f}^i\phi$$

全ての感情についての証明は省略し、*Joy* と *Satisfaction* の2つの感情について証明する。

#### 証明

$\vdash Joy_{f_J(d)}^i\phi \rightarrow Bel^iJoy_{f_J(d)}^i\phi$  を示す。

*Joy* の定義より、 $Joy_{f_J(d)}^i\phi := Bel^i\phi \wedge Des_d^i\phi$  なので、

$Bel^i\phi \wedge Des_d^i\phi \rightarrow Bel^i(Bel^i\phi \wedge Des_d^i\phi)$  を証明する。

ある世界  $w$  で  $Bel^i\phi \wedge Des_d^i\phi$  が真であると仮定すると、

$w$  で  $Bel^i\phi$ 、 $Des_d^i\phi$  が真である。

(4-Bel<sup>i</sup>) と (4-MIX) より、  
 $w$  で  $Bel^i Bel^i \phi$ 、 $Bel^i Des^i_d \phi$  が真となる。  
(C-□) より、 $w$  で  $Bel^i (Bel^i \phi \wedge Des^i_d \phi)$  が真となる。  
以上より、世界  $w$  で  $Bel^i \phi \wedge Des^i_d \phi \rightarrow Bel^i (Bel^i \phi \wedge Des^i_d \phi)$  が成り立つ。  
つまり、 $Joy^i_{f_J(d)} \phi \rightarrow Bel^i Joy^i_{f_J(d)} \phi$  が成り立つ。  
 $\vdash Bel^i Joy^i_{f_J(d)} \phi \rightarrow Joy^i_{f_J(d)} \phi$  についても、同様に示すことができる。

## 証明

$\vdash Satisfaction^i_{f_{St}(x,e,\ell)} \phi \rightarrow Bel^i Satisfaction^i_{f_{St}(x,e,\ell)} \phi$  を示す。

$\ell$  を定数とする。

Satisfaction の定義より、

$Satisfaction^i_{f_{St}(x,e,\ell)} := Bel^i PHope^i_x \phi \wedge Bel^i PEffort^i_e \phi \wedge Prob^i_\ell \phi$  なので、  
 $Bel^i PHope^i_x \phi \wedge Bel^i PEffort^i_e \phi \wedge Prob^i_\ell \phi \rightarrow Bel^i (Bel^i PHope^i_x \phi \wedge Bel^i PEffort^i_e \phi \wedge Prob^i_\ell \phi)$   
を証明する。

ある世界  $w$  で、 $Bel^i PHope^i_x \phi \wedge Bel^i PEffort^i_e \phi \wedge Prob^i_\ell \phi$  が真であると仮定すると、  
 $w$  で  $Bel^i PHope^i_x \phi$ 、 $Bel^i PEffort^i_e \phi$ 、 $Prob^i_\ell \phi$  が真である。

(4-Bel<sup>i</sup>) と (4-MIX) より、

$w$  で  $Bel^i Bel^i PHope^i_x \phi$ 、 $Bel^i Bel^i PEffort^i_e \phi$ 、 $Bel^i Prob^i_\ell \phi$  が真となる。

(C-□) より、 $w$  で  $Bel^i (Bel^i PHope^i_x \phi \wedge Bel^i PEffort^i_e \phi \wedge Prob^i_\ell \phi)$  が真となる。

以上より、世界  $w$  で

$Bel^i PHope^i_x \phi \wedge Bel^i PEffort^i_e \phi \wedge Prob^i_\ell \phi \rightarrow Bel^i (Bel^i PHope^i_x \phi \wedge Bel^i PEffort^i_e \phi \wedge Prob^i_\ell \phi)$   
が成り立つ。

つまり、 $Satisfaction^i_{f_{St}(x,e,\ell)} \phi \rightarrow Bel^i Satisfaction^i_{f_{St}(x,e,\ell)} \phi$  が成り立つ。

$\vdash Bel^i Satisfaction^i_{f_{St}(x,e,\ell)} \phi \rightarrow Satisfaction^i_{f_{St}(x,e,\ell)} \phi$  についても同様に示すことができる。