

A Formal Notation for the Elicitation of Emotions based on the OCC Model

Jean-Paul Sansonnet and François Bouchet¹

Abstract. The purpose of this paper is to explore the OCC model as described in [3], called OCC₈₈ in the following, and the more recent version presented in [2], further called OCC₀₂. On this basis, we propose a new classification of emotions based on the classes of procedural situations in which the emotions are elicited. First we present a formalization of the OCC₈₈ model taken from [4]; then we propose a first transformation based on the formalization of the activities. In a second study we present a formalization of the OCC₀₂ model and which is again classified in terms of activities. Finally both the OCC₈₈ and OCC₀₂ models of emotion are classified in a new taxonomy which focuses on the nature of the arguments of the emotions, hence revealing *holes* in the OCC taxonomy which are filled with new emotions.

1 THE OCC₈₈ MODEL

The first OCC model was proposed by Ortony et al [3] in 1988. This model describes a hierarchy that classifies 22 emotions of human beings, further called 'agents'. Actually, they are defined as 11 emotions composed of a positive pole and a negative pole (e.g. love/hate, pride/shame...). In this model, referred to as OCC₈₈, emotions are not intrinsic to the agent (like in a pure mental state of "fear") but instead are relative to some entities, external to the agent, that *trigger* the emotion (Ortony also says that emotions are *elicited*) in the mental state of the agent. These entities are divided into three classes :

- **Objects** of the world, e.g. other agents or physical entities,
- **Actions** (and plans) performed by agents in the world,
- **Consequences of events** in the world, either achieved or considered as potential (future).

Moreover, an extra class is "compound emotions" that group consequences of events caused by actions performed by agents (*cf.* below gratitude/anger).

1.1 A formal model for the OCC₈₈ model

In order to discuss formally the OCC₈₈ model, we use here the formal representation proposed in [4]. We define the following elements² :

$\mathcal{G} = \{i, j, \dots\}$	the set of agents
$\mathcal{A} = \{\alpha_1, \alpha_2, \dots\}$	the set of actions
$\Pi = \{\pi_1, \pi_2, \dots\}$	the set of plans
$\mathcal{K} = \{\kappa_1, \kappa_2, \dots\}$	the set of states of the world

¹ LIMSI-CNRS, BP 133 F-91403 Orsay cedex, France. Email: {jps,bouchet} at limsi.fr

² In the formalization proposed by [4], objects are restricted to agents.

A plan $\pi \in \Pi$ is a sequence of actions, such as $\pi = \langle \alpha_1; \alpha_2; \alpha_3 \rangle$. A state $\kappa \in \mathcal{K}$ is a conjunction of literals (atomic propositions or their negation). If κ represents a goal to reach, that goal can thus be broken up into the set of accomplished subgoals (κ^+) and the set of remaining subgoals (κ^-). Since we require goals to be logically consistent and non empty, we do not consider the empty conjunction.

Moreover, we define \mathcal{E} as the set of the 22 emotions from the OCC₈₈ vocabulary, such as :

$\mathcal{E} = \{\text{gratification, remorse, gratitude, anger, pride, shame, admiration, reproach, joy, distress, happy-for, resentment, gloating, pity, hope, fear, satisfaction, disappointment, relief, fears-confirmed, love, hate}\}.$

We write ε_i an emotion $\varepsilon \in \mathcal{E}$ felt by the agent i . For each emotion, it is possible to associate a list of elements (x, y, \dots) that have an impact in the triggering of that emotion, which will be written as $\varepsilon_i(x, y, \dots)$. The table 1 displays this information for the 22 emotions of OCC₈₈.

Table 1. Emotion triggering fluents for OCC₈₈ emotions, according to [4]

Positive emotions	Negative emotions
gratification _i (α, κ)	remorse _i (α, κ)
gratitude _i (j, α, κ)	anger _i (j, α, κ)
pride _i (α)	shame _i (α)
admiration _i (j, α)	reproach _i (j, α)
joy _i (κ)	distress _i (κ)
happy-for _i (j, κ)	resentment _i (j, κ)
gloating _i (j, κ)	pity _i (j, κ)
hope _i (π, κ)	fear _i (π, κ)
satisfaction _i (π, κ)	disappointment _i (π, κ)
relief _i ($\pi, \neg\kappa$)	fears-confirmed _i ($\pi, \neg\kappa$)
love _i (j)	hate _i (j)

1.2 Transformation of the notation

We can classify the 11 bivalent OCC₈₈ emotions according to the structure of the list of triggering arguments, distinguishing:

- Mono-arguments: α, π, i, κ
- Bi-arguments: $(\alpha, \kappa), (j, \alpha), (j, \kappa), (\pi, \kappa)$
- Tri-arguments (compound emotions): (j, α, κ)

Among those, we can distinguish two main cases:

1. intrinsic emotional reactions: when the only agent involved is the agent i itself,

2. relational emotional reactions: when there is agent i and a second agent j involved³.

We define the following notations:

$arg \mapsto \pm\varepsilon$	the argument triggers a positive/negative emotion
$\kappa(i)$	agent i experiences the currently holding state κ
$i(\alpha)$	agent i executes action α
$i(\alpha \pi)$	agent i executes action α or plan π
$exec \Rightarrow \kappa$	the execution entailed state κ to hold
$exec \stackrel{?}{\Rightarrow} \kappa$	the execution may entail state κ to hold in the future
$i \prec arg$	agent i reacts emotionally to a considered argument

With these notations, we can rewrite the definition of the OCC₈₈ emotions according to table 2.

1.3 Remarks

With these formal notations of the 22 OCC₈₈ emotions, we can make some remarks:

- Plans are introduced as sequences of actions, but partially achieved plans are not used to define the emotions so we can deal only with plans and consider actions as singletons: $\alpha_1 = \pi_1 = \langle \alpha_1 \rangle$.
- The OCC₈₈ model deals mainly with *achieved* consequences of agents' actions/plans. The exception is the fear/hope emotion couple where consequences of agents' actions/plans are *envisioned* as potential steps. This is probably due to the fact that the OCC₈₈ model considers the *triggering* of emotions from *achieved* events in the world.
- When we consider the expression $i(\alpha) \mapsto \pm\varepsilon$ which triggers shame/pride, a question arises about the actual object that triggers the emotion: a) as defined in [4], it can be the action itself α_1 , or, alternatively, it can be the state κ in which the agent stands after having executed the action done (α_1). One can consider that these options are either equivalent or not significant; then it is possible to rewrite the expression as $(i(\alpha) \Rightarrow \kappa) \mapsto \pm\varepsilon$.
- Now if we consider for example the emotion $\varepsilon = anger$, we can note that the agent i experiences this emotion in three ways:
 - $anger_i()$ means that agent i experiences pure anger;
 - $anger_i(x)$ means that agent i experiences anger *about* an object x (agent, thing, concept);
 - $anger_i(x, r)$ means that agent i experiences anger about x for the *reason* r (cause, consequence).

One can debate whether pure anger is relevant beyond expressing a pure bodily arousal without any ongoing cognitive process. In that case, emotions are defined in terms of two arguments: an object and a reason.

2 THE OCC₀₂ MODEL

2.1 Presentation

The OCC₀₂ model was proposed by Ortony in [2]. This classification is quite informal but is more systematic and fixes some of the problems of the first version, as stated by Ortony himself: "... I think we might want to consolidate some of our categories of emotions. So, instead of the rather cumbersome (and to some degree arbitrary)

analysis we proposed in 1988, I think it worth considering collapsing some of the original categories down to five distinct positive and five negative specializations of two basic types of affective reactions – positive and negative ones ...". That updated classification is shown in table 3, where the first entry in each group of six is the undifferentiated (positive or negative) reaction, and the remaining five entries are specializations: the first pair (2-3) is goal-based, the second (4-5) is standards-based and the last (6) is taste-based.

2.2 A formal model for the OCC₀₂ model

2.2.1 Notations

Transitions: We define a transition as a 3-tuple $\tau = \langle \pi, i, \kappa \rangle$ where $\pi \in \Pi$, $i \in \mathcal{G}$, $\kappa \in \mathcal{K}$ and we use the notation $\tau : \pi_i \Rightarrow \kappa$ which means that the agent i executes a plan (or action) π resulting in state κ to hold.

We distinguish two situations, according to the position of the transition τ from the point of view of the agent i :

- $\tau/i \mapsto \varepsilon_i$ means that the transition τ has been *achieved*, and following that execution, the agent i experiences the emotion ε ,
- $\tau \setminus i \mapsto \varepsilon_i$ means that the agent i *envisions* the possibility that transition might be executed in the future and, henceforce, experiences the emotion ε .

Plans: The same way we note π_i a plan executed by the agent i , we write π_* a plan executed by any agent.

States: We note κ_{j+} (resp. κ_{j-}) the fact that the state κ holds and is evaluated (by i) as good (resp. bad) for j . We note $\neg\kappa$ the fact that the state κ doesn't hold.

Emotions: We note:

- ε_i the fact that emotion ε is experienced by the agent i ,
- $\varepsilon_i(a)$ the fact that emotion ε is experienced by the agent i and refers to the argument a ,
- $\varepsilon|\varepsilon'$ an alternative between two emotions ε and ε' .

2.2.2 Formal transcription of OCC₀₂

The six cases are described in our formal notation in table 4, each of the 12 lines corresponding to one from the table 3 (written as $[number]^{\text{valence}}$, e.g. 4^+ for "a self-initiated praiseworthy act"). We also associate to them the 22 emotion names taken from OCC₈₈.

We see that the use of that notation to describe situations where emotions shall be triggered can't be completely matched with the emotions from OCC₈₈ or OCC₀₂. For instance, compared to OCC₈₈ formalization, we see that anger ends up appearing in two situations (as $anger_i$ and $anger_i(j)$), admiration is not directly opposed to reproach, gloating and pity are together as a possible reaction to a same situation, etc.

2.3 A new ontology of emotions

2.3.1 Argument-oriented redefinition of emotions

Moreover, concerning $\pi_{j \neq i}$ situations, the agent i experiences emotion ε_i about the other agent j as the priority argument. This raises a question about the arguments of the other emotions, and thus lead us to propose a new redefined set of emotions, according to their

³ In [4], at most two agents i and j can be involved.

Table 2. Rewritten OCC₈₈ emotions for an agent i

Intrinsic reactions			Relational reactions ($i \neq j$)		
Description	Emotions		Description	Emotions	
	-	+		-	+
	-	-		-	-
$i(\alpha) \mapsto \pm\varepsilon$	shame _{i}	pride _{i}	$i \prec \cdot j \mapsto \pm\varepsilon$	hate _{i}	love _{i}
$\kappa(i) \mapsto \pm\varepsilon$	distress _{i}	joy _{i}	$i \prec \cdot j(\alpha) \mapsto \pm\varepsilon$	reproach _{i}	admiration _{i}
			$i \prec \cdot \kappa(j) \mapsto \pm\varepsilon$	resentment _{i}	happy-for _{i}
$(i(\alpha) \Rightarrow \kappa) \mapsto \pm\varepsilon$	remorse _{i}	gratification _{i}	$(i \prec \cdot j(\alpha) \Rightarrow \kappa) \mapsto \pm\varepsilon$	pity _{i}	gloating _{i}
$(i(\pi) \Rightarrow \kappa) \mapsto \pm\varepsilon$	disappointment _{i}	satisfaction _{i}		anger _{i}	gratitude _{i}
$(i(\pi) \Rightarrow \neg\kappa) \mapsto \pm\varepsilon$	fears-confirmed _{i}	relief _{i}			
$(i(\alpha \pi) \stackrel{?}{\Rightarrow} \kappa) \mapsto \pm\varepsilon$	fear _{i}	hope _{i}		-	-

Table 3. Five specializations of generalized good and bad feelings (by [2], collapsed from [3])

N ^o	Positive reactions		Negative reactions	
	Cause	Examples	Cause	Examples
1	something good happened	joy, happiness	something bad happened	distress, sadness
2	the possibility of something good happening	hope	the possibility of something bad happening	fear
3	a feared bad thing didn't happen	relief	a hoped-for good thing didn't happen	disappointment
4	a self-initiated praiseworthy act	pride, gratification	a self-initiated blameworthy act	remorse, self-anger, shame
5	an other-initiated praiseworthy act	gratitude, admiration	an other-initiated blameworthy act	anger, reproach
6	someone/thing is found appealing/attractive	love, like	someone/thing is found unappealing/unattractive	hate, dislike

Table 4. Formal representation of situations and corresponding emotion and cases in OCC₈₈ and OCC₀₂

Formal notation of a situation	Emotion in OCC ₈₈	Case in OCC ₀₂
$\pi_* \Rightarrow \kappa_{i+}/i \mapsto$	joy _{i}	1 ⁺
$\pi_* \Rightarrow \kappa_{i-}/i \mapsto$	distress _{i}	1 ⁻
$\pi_i \Rightarrow \kappa_{i+}/i \mapsto$	pride _{i} satisfaction _{i} gratification _{i}	4 ⁺
$\pi_i \Rightarrow \kappa_{i-}/i \mapsto$	anger _{i} fears-confirmed _{i} shame _{i} remorse _{i}	4 ⁻
$\pi_i \Rightarrow \neg(\kappa_{i+})/i \mapsto$	disappointment _{i}	3 ⁻
$\pi_i \Rightarrow \neg(\kappa_{i-})/i \mapsto$	relief _{i}	3 ⁺
$\pi_{j \neq i} \Rightarrow \kappa_{i+}/i \mapsto$	gratitude _{i} (j)	5 ⁺
$\pi_{j \neq i} \Rightarrow \kappa_{i-}/i \mapsto$	anger _{i} (j) reproach _{i} (j)	5 ⁻
$\pi_{j \neq i} \Rightarrow \kappa_{j+}/i \mapsto$	happy-for _{i} (j) admiration _{i} (j) resent _{i} (j)	5 ⁺
$\pi_{j \neq i} \Rightarrow \kappa_{j-}/i \mapsto$	pity _{i} (j) gloating _{i} (j)	5 ⁻
$\pi_* \Rightarrow \kappa_{i+}/i \mapsto$	hope _{i}	2 ⁺
$\pi_* \Rightarrow \kappa_{i-}/i \mapsto$	fear _{i}	2 ⁺
$x/i \mapsto$	love _{i} (x) hate _{i} (x)	6 ⁺ 6 ⁻

arguments. We now use the notation $\varepsilon_i(x)$ to mean that agent i experiences emotion ε “directed in priority towards” the argument x . In the list below, emotions written in bold are created to fill the gaps appearing in OCC₈₈ and OCC₀₂ with this new way to consider the emotions. We thus have :

1. emotions with $x = \text{object} \mid (\text{agent } j \neq i) \mid \text{concept}$:

+ love _{<i>i</i>} (x)	– hate _{<i>i</i>} (x)
------------------------------------	------------------------------------
2. emotions with $x = (\text{agent } j \neq i)$:

+ admiration _{<i>i</i>} (j)	– despect _{<i>i</i>} (j)
+ gratefulness _{<i>i</i>} (j)	– ungratefulness _{<i>i</i>} (j)
+ friendliness _{<i>i</i>} (j)	– anger _{<i>i</i>} (j)
+ congratulation _{<i>i</i>} (j)	– reproach _{<i>i</i>} (j)
+ happiness-for _{<i>i</i>} (j)	– resentment _{<i>i</i>} (j)
+ gloating _{<i>i</i>} (j)	– pity _{<i>i</i>} (j)
3. emotions with $x = \text{agent } i$:

+ admiration _{<i>i</i>} (i)	– despect _{<i>i</i>} (i)
→ pride _{<i>i</i>}	→ shame _{<i>i</i>}
+ gratefulness _{<i>i</i>} (i)	– ungratefulness _{<i>i</i>} (i)
→ \emptyset	→ \emptyset
+ friendliness _{<i>i</i>} (i)	– anger _{<i>i</i>} (i)
→ \emptyset	→ self-anger _{<i>i</i>}
+ congratulation _{<i>i</i>} (i)	– reproach _{<i>i</i>} (i)
→ gratification _{<i>i</i>}	→ remorse _{<i>i</i>}
+ happiness-for _{<i>i</i>} (i)	– resentment _{<i>i</i>} (i)
→ self-satisfaction _{<i>i</i>}	→ \emptyset
+ gloating _{<i>i</i>} (i)	– pity _{<i>i</i>} (i)
→ self-deprecation _{<i>i</i>}	→ self-pity _{<i>i</i>}
4. emotions with $x = \tau \mid \neg\tau$:

+ hope _{<i>i</i>} (τ)	– fear _{<i>i</i>} (τ)
+ relief _{<i>i</i>} ($\neg\tau$)	– fears-confirmed _{<i>i</i>} (τ)
+ expectation _{<i>i</i>} (τ)	– surprise _{<i>i</i>} (τ)
5. emotions with $x = \emptyset \mid \Omega$ ($\Omega = \text{everything}$):

+ joy _{<i>i</i>} (\emptyset)	– distress _{<i>i</i>} (\emptyset)
---	--

We can make a few remarks about that list :

- Here emotions are considered as mental attitudes directed towards an object (an argument), which makes that joy and distress can be viewed as attitudes towards nothing in particular.
- Among the added emotions, we can mention the surprise and the despise (judging something shameful or disgusting), which are parts of the five primary emotions considered by Ekman [1] as universal in terms of facial expression. The three other ones (joy, anger and fear) were already in the OCC₈₈ list.
- Expectation and surprise are opposed but can’t be really given a valence like the other emotions.
- Disappointment is not the equivalent of resentment of i towards itself and is the only one of the 22 emotions from OCC₈₈ that doesn’t appear at all here (and should maybe be added).

2.3.2 Classes of mental attitudes towards an object

From the previous list, we can, as Ortony did between OCC₈₈ and OCC₀₂, try to collapse our list of emotions into main classes of mental attitudes. We thus define 7 classes of mental attitudes towards an object (written in upper case):

- | | | |
|---|---|----------|
| + HAPPY _{<i>i</i>} (\emptyset) | – SAD _{<i>i</i>} (\emptyset) | 1 agent |
| + LIKE _{<i>i</i>} (x) | – DISLIKE _{<i>i</i>} (x) | 2 agents |
| friendly _{<i>i</i>} (j) | angry _{<i>i</i>} ($i \mid j$) | |
| love _{<i>i</i>} ($i \mid j$) | hate _{<i>i</i>} ($i \mid j$) | |
| + SYMPATHY _{<i>i</i>} (x) | – ANTIPATHY _{<i>i</i>} (x) | 2 agents |
| happy-for _{<i>i</i>} (j) | resent _{<i>i</i>} (j) | |
| self-satisfaction _{<i>i</i>} | \emptyset | |
| pity _{<i>i</i>} (j) | gloating _{<i>i</i>} (j) | |
| self-deprecation _{<i>i</i>} | \emptyset | |
| self-pity _{<i>i</i>} | \emptyset | |
| + PRAISE _{<i>i</i>} (x) | – DESPISE _{<i>i</i>} (x) | 2 agents |
| admiration _{<i>i</i>} (j) | reproach _{<i>i</i>} (j) | |
| pride _{<i>i</i>} | shame _{<i>i</i>} | |
| gratification _{<i>i</i>} | remorse _{<i>i</i>} | |
| + THANKFUL _{<i>i</i>} (x) | – UNGRATEFUL _{<i>i</i>} (x) | 2 agents |
| + HOPE _{<i>i</i>} (x) | – FEAR _{<i>i</i>} (x) | 1 thing |
| relief _{<i>i</i>} (x) | fears-confirmed _{<i>i</i>} (x) | |
| + EXPECT _{<i>i</i>} (x) | – SURPRISE _{<i>i</i>} (x) | 1 thing |

Note : the name of emotions aren’t exactly corresponding to the ones in the previous list.

3 CONCLUSION

In this work we have explored the OCC₈₈ and OCC₀₂ models from the point of view of the procedural situation in which the emotions are elicited. then we have made another classification which focuses on the nature of the arguments of the emotions. this has resulting in a proposition of a taxonomy of emotions based on 7 groups of bivalents emotions.

REFERENCES

- [1] Paul Ekman, Wallace V. Friesen, and Phoebe C. Ellsworth, Emotion in the human face, Pergamon Press, May 1972.
- [2] Andrew Ortony, ‘On making believable emotional agents believable’, in Emotions in humans and artifacts, eds., Robert Trappl, Paolo Petta, and Sabine Payr, 189–211, MIT Press, Cambridge, MA, (2003).
- [3] De Andrew Ortony, Gerald L. Clore, and Allan Collins, The Cognitive Structure of Emotions, Cambridge, UK, cambridge university press edn., 1988.
- [4] Bas R. Steunebrink, Mehdi Dastani, and John-Jules Ch. Meyer, ‘A formal model of emotions: Integrating qualitative and quantitative aspects’, in Proceeding of the 2008 conference on ECAI 2008: 18th European Conference on Artificial Intelligence, pp. 256–260. IOS Press, (2008).