A logic of emotions: from appraisal to coping

Mehdi Dastani
Utrecht University, The Netherlands
mehdi@cs.uu.nl

Emiliano Lorini
IRIT-CNRS, Toulouse, France
lorini@irit.fr

ABSTRACT

Emotion is a cognitive mechanism that directs an agent's thoughts and attention to what is relevant, important, and significant. Such a mechanism is crucial for the design of resource-bounded agents that must operate in highly-dynamic, semi-predictable environments and which need mechanisms for allocating their computational resources efficiently. The aim of this work is to propose a logical analysis of emotions and their influences on an agent's behavior. We focus on four emotion types (*viz.*, hope, fear, joy, and distress) and provide their logical characterizations in a modal logic framework. As the intensity of emotion is essential for its influence on an agent's behavior, the logic is devised to represent and reason about graded beliefs, graded goals and intentions. The belief strength and the goal strength determine the intensity of emotions. Emotions trigger different types of coping strategy which are aimed at dealing with emotions either by forming or revising an intention to act in the world, or by changing the agent's interpretation of the situation (by changing beliefs or goals).

Categories and Subject Descriptors

I.2 [**Artificial Intelligence**]: Intelligent agents

General Terms

Theory

Keywords

Cognitive models, logic-based approaches and methods

1. INTRODUCTION

Autonomous software agents are assumed to have different (possibly conflicting) objectives, able to sense their environments, update their states accordingly, and decide which actions to perform at any moment in time. The behavior of such software agents can be effective and practical only if they are able to continuously and adequately assess their (sensed) situation and update their states with relevant information and crucial objectives. For example, a robot with a plan to transport a container to its target position may perceive it has low battery charge. The robot may assess the state of its battery charge as being relevant for the objective of having the container at its target position, and update its state by suspending the current battery-demanding transport plan. Such assessment and update may cause the agent to decide to charge its battery right away or to focus on a less battery-demanding task.

Emotion is a (cognitive) mechanism that directs one's thoughts and attention to what is relevant, important, and significant in order to ensure effective behavior. The aim of this work is to propose a logical analysis of the relationships between emotion and cognition. An understanding of these relationships is particularly important in the perspective of the design of resource-bounded agents that must survive in highly-dynamic, semi-predictable environments and which need mechanisms for allocating their computational resources efficiently. Indeed, as it has been stressed by several authors in psychology and economics, emotions provide heuristics for preventing excessive evaluation and deliberation by pruning of search spaces [5] and for interrupting normal cognition when unattended goals require servicing [20], and signals for belief revision [15].

Our approach is inspired by the appraisal and coping models of human emotions [16, 11, 8]. According to these models, an agent continuously appraises its situation (*e.g.*, low battery charge endangers the objective of having a container at its target position) after which emotions can be triggered (*e.g.*, fear of failing to place the container at its target position). The triggered emotions can affect the agent's behavior depending on their intensities. There are often a set of strategies that can be used to cope with a specific emotion, for example, by updating the agent's mental state (*e.g.*, being fearful that the transportation plan will not place the container at its target position leads the agent to reconsider its plan).

We first propose, in Section 2, a dynamic logic with special operators which allow to represent the intentions of a cognitive agent as well as its beliefs and goals with their corresponding strengths. Then, in Section 3, we provide a logical analysis of the intensity of different emotions such as hope, fear, joy and distress. In Section 4, we extend the logic with special operators to formally characterize different kinds of coping strategies which are aimed at dealing with emotions either by forming or revising an intention to act in the world, or by changing the agent's interpretation of the situation (by changing belief strength or goal strength). A complete axiomatization and a decidability result for the logic are given in Section 5. Related works are discussed in Section 6.

2. LOGICAL FRAMEWORK

This section presents the syntax and the semantics of the logic DL-GA (*Dynamic Logic of Graded Attitudes*). This logic is designed to represent beliefs, goals, and intentions, where beliefs and goals have degree of plausibility and desirability, respectively.

2.1 Syntax

Assume a finite set of atomic propositions describing facts $Atm = \{p, q, \ldots\}$, a finite set of physical actions (*i.e.*, actions modifying the physical world) $PAct = \{a, b, \ldots\}$, a finite set of natural numbers

Appears in: *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012)*, Conitzer, Winikoff, Padgham, and van der Hoek (eds.), 4-8 June 2012, Valencia, Spain.

$Num = \{x \in \mathbb{N} : 0 \leq x \leq \mathsf{max}\}$, with $\mathsf{max} \in \mathbb{N} \setminus \{0\}$. We note $Num^- = \{-x : x \in Num \setminus \{0\}\}$ the corresponding set of negative integers. We note *Lit* the set of literals and $l, l', \ldots$ the elements of *Lit*. Finally, we define the set of propositional formulas *Prop* to be the set of all Boolean combinations of atomic propositions.

The language $\mathcal{L}$ of DL-GA is defined by the following grammar in Backus-Naur Form (BNF):

$$\begin{array}{lllll} Act & : & \alpha & ::= & a \mid *\varphi \\ Fml & : & \varphi & ::= & p \mid \mathsf{exc_h} \mid \mathsf{Des^k} l \mid \mathsf{Int}_a \mid \\ & & & & \neg\varphi \mid \varphi \wedge \varphi \mid \mathsf{K}\varphi \mid [\alpha]\varphi \end{array}$$

where p ranges over *Atm*, h ranges over *Num*, l ranges over *Lit*, k ranges over $Num \cup Num^-$, and a ranges over *PAct*. The other Boolean constructions $\top$, $\bot$, $\vee$, $\rightarrow$ and $\leftrightarrow$ are defined in the standard way.

The set of actions *Act* includes both physical actions and sensing actions of the form $*\varphi$. A sensing action is an action which consists in modifying the agent's beliefs in the light of a new incoming evidence. In particular, $*\varphi$ is the mental action (or process) of learning that φ is true. As we will show in Section 2.3, technically this amounts to an operation of belief conditioning in Spohn's sense [21].

The set of formulas *Fml* contains special constructions $\mathsf{exc_h}$, $\mathsf{Des^k} l$ and Int_a which are used to represent the agent's mental state. Formulas $\mathsf{exc_h}$ are used to identify the degree of *plausibility* of a given world for the agent. We here use the notion of plausibility first introduced by Spohn [21]. Following Spohn's theory, the worlds that are assigned the smallest numbers are the most plausible, according to the beliefs of the individual. That is, the number h assigned to a given world rather captures the degree of *exceptionality* of this world, where the exceptionality degree of a world is nothing but the opposite of its plausibility degree (*i.e.*, the exceptionality degree of a world decreases when its plausibility degree increases). Therefore, formula $\mathsf{exc_h}$ can be read alternatively as "the current world has a degree of exceptionality h" or "the current world has a degree of plausibility $\mathsf{max} - \mathsf{h}$".

Formula $\mathsf{Des^k} l$ has to be read "the state of affairs l has a degree of desirability k for the agent". Degree of desirability can be positive, negative or equal to zero.[1] Suppose $\mathsf{k} > 0$. Then $\mathsf{Des^k} l$ means that "the agent wishes to achieve l with strength k", whereas $\mathsf{Des^{-k}} l$ means that "the agent wishes to avoid l with strength k". $\mathsf{Des^0} l$ means that "the agent is indifferent about l" (*i.e.*, the agent does not care whether l is true or false). For notational convenience, in what follows we will use the following abbreviations:

$$\mathsf{AchG^k} l \stackrel{\text{def}}{=} \mathsf{Des^k} l \ \text{ for } \mathsf{k} > 0$$
$$\mathsf{AvdG^k} l \stackrel{\text{def}}{=} \mathsf{Des^{-k}} l \ \text{ for } \mathsf{k} > 0$$

where AchG andAvdG respectively stand for *achievement goal* and *avoidance goal*.

Formulas Int_a capture the agent's intentions. We assume that the agent's intentions are only about physical actions and not about sensing actions. The formula Int_a has to be read "the agent has the intention to perform the physical action a" or "the agent is committed to perform the physical action a".

The logic DL-GA has also epistemic operators and modal operators that are used to describe the effects of a given action α. The formula $[\alpha]\varphi$ has to be read "after the occurrence of the action α, φ will be true". $\mathsf{K}\varphi$ has to be read "the agent knows that φ is true". This concept of knowledge is the standard S5-notion, partition-based and fully introspective, that is commonly used in computer science and economics [7]. The operator $\widehat{\mathsf{K}}$ is the dual of K, that is, $\widehat{\mathsf{K}}\varphi \stackrel{\text{def}}{=} \neg\mathsf{K}\neg\varphi$. As we will show in the Section 2.5, the operator K captures a form of 'absolutely unrevisable belief', that is, a form of belief which is stable under belief revision with any new *evidence*.

[1]However, we assume that exceptionality/plausibility and positive desirability are measured on the same scale *Num*.

2.2 Physical action description

Similarly to Situation Calculus [18], in our framework physical actions are described in terms of their executability preconditions and of their positive and negative effect preconditions. In particular, we define a function

$$Pre : PAct \longrightarrow Prop$$

to map physical actions to their executability preconditions. Using the notion of executability precondition, we define special dynamic operators for physical actions of the form $\langle\langle a\rangle\rangle$, where $\langle\langle a\rangle\rangle\varphi$ has to be read " the physical action a is executable and, φ will be true afterwards":

$$\langle\langle a\rangle\rangle\varphi \stackrel{\text{def}}{=} Pre(a) \wedge [a]\varphi$$

Moreover, we introduce two functions

$$\gamma^+ : PAct \times Atm \longrightarrow Prop$$
$$\gamma^- : PAct \times Atm \longrightarrow Prop$$

mapping physical actions and atomic propositions to propositional formulas. The formula $\gamma^+(a, p)$ describes the *positive effect preconditions* of action a with respect to p, whereas $\gamma^-(a, p)$ describes the *negative effect preconditions* of action a with respect to p. The former represent the necessary and sufficient conditions for ensuring that p will be true after the occurrence of the physical action a, while the latter represent the necessary and sufficient conditions for ensuring that p will be false after the occurrence of the physical action a. We make the following *coherence assumption*:

(COH_γ) for every $a \in PAct$ and $p \in Atm$, $\gamma^+(a, p)$ and $\gamma^-(a, p)$ must be logically inconsistent.

COH_γ ensures that actions do not have contradictory effects.

2.3 Models and truth conditions

The semantics of the logic DL-GA is a possible world semantics with special functions for exceptionality, desirability and intentions.

Definition 1 (Model). *DL-GA models are tuples $M = \langle W, \sim, \kappa_{\mathsf{exc}}, \mathcal{D}, \mathcal{I}, \mathcal{V}\rangle$ where:*

- *W is a nonempty set of possible worlds or states;*
- *$\sim$ is an equivalence relation between worlds in W;*
- *$\kappa_{\mathsf{exc}} : W \longrightarrow Num$ is a total function from the set of possible worlds to the set of natural numbers Num;*
- *$\mathcal{D} : W \times Lit \longrightarrow Num \cup Num^-$ is a total function from the set of possible worlds to the set of integers $Num \cup Num^-$;*
- *$\mathcal{I} : W \longrightarrow 2^{PAct}$ is a total function called* commitment function*, mapping worlds to sets of physical actions;*
- *$\mathcal{V} : W \longrightarrow 2^{Atm}$ is a valuation function.*

As usual, $p \in \mathcal{V}(w)$ means that proposition p is true at world w. The equivalence relation $\sim$, which is used to interpret the epistemic operator K, can be viewed as a function from W to 2^W. Therefore, we can write $\sim(w) = \{v \in W : w \sim v\}$. The set $\sim(w)$ is the agent's *information state* at world w: the set of worlds that the agent imagines at world w. As $\sim$ is an equivalence relation, if $w \sim v$ then the agent has the same information state at w and v (*i.e.*, the agent imagines the same worlds at w and v).

The function κ_{exc} represents a plausibility grading of the possible worlds and is used to interpret the atomic formulas $\mathsf{exc_h}$.

$\kappa_{\mathsf{exc}}(w) = \mathsf{h}$ means that, according to the agent the world w has a degree of exceptionality h or, alternatively, according to the agent the world w has a degree of plausibility $\mathsf{max} - \mathsf{h}$. (Remember that the degree of plausibility of a world is the opposite of its exceptionality degree). The function κ_{exc} allows to model the notion of belief: among the worlds the agent cannot distinguish from a given world w (*i.e.*, the agent's information state at w), there are worlds that the agent considers more plausible than others. For example, suppose that $\sim(w) = \{w, v, u\}$, $\kappa_{\mathsf{exc}}(w) = 2$, $\kappa_{\mathsf{exc}}(u) = 1$ and $\kappa_{\mathsf{exc}}(v) = 0$. This means that $\{w, v, u\}$ is the set of worlds that the agent imagines at world w. Moreover, according to the agent, the world v is strictly more plausible than the world u and the world u is strictly more plausible than the world w (as $\mathsf{max} - 0 > \mathsf{max} - 1 > \mathsf{max} - 2$).

DL-GA models are supposed to satisfy the following *normality* constraint for the plausibility grading to ensure that the agent's beliefs are consistent:

($Norm_{\kappa_{\mathsf{exc}}}$) for every $w \in W$, there is v such that $w \sim v$ and $\kappa_{\mathsf{exc}}(v) = 0$.

The function $\mathcal{D}$ is used to interpret the atomic formulas $\mathsf{Des}^{\mathsf{k}} l$. Suppose $\mathsf{k} > 0$. Then, $\mathcal{D}(w, l) = \mathsf{k}$ means that, at world w, l has a degree of desirability k; whereas $\mathcal{D}(w, l) = -\mathsf{k}$ means that, at world w, l has a degree of desirability $-\mathsf{k}$ — or equivalently, l has a degree of undesirability k —. $\mathcal{D}(w, l) = 0$ means that the agent is indifferent about l.

The function $\mathcal{I}$ is used to interpret the atomic formulas Int_a. For every world $w \in W$, $\mathcal{I}(w)$ identifies the set of physical actions that the agent intends to perform. $\mathcal{I}(w) = \emptyset$ means that the agent has no intention.

Note that in our dynamic setting an agent may be committed to perform an action even though it believes that this is a *suboptimal* choice, *i.e.*, we do not require agents to have intentions because of their desirable consequences. An agent may have an intention without desiring its consequence because, for example, its beliefs and desires may change due to a sensing action. In our running example, the robot may have the intention to transport a container to a given target position, while it believes that this is a suboptimal choice as it has just learnt that it does not have sufficient battery power to accomplish the task.

Definition 2 (Truth conditions). *Given a* DL-GA *model M, a world w and a formula φ, $M, w \models \varphi$ means that φ is true at world w in M. The rules defining the truth conditions of formulas are:*

- $M, w \models p$ *iff* $p \in \mathcal{V}(w)$
- $M, w \models \mathsf{exc}_{\mathsf{h}}$ *iff* $\kappa_{\mathsf{exc}}(w) = \mathsf{h}$
- $M, w \models \mathsf{Des}^{\mathsf{k}} l$ *iff* $\mathcal{D}(w, l) = \mathsf{k}$
- $M, w \models \mathsf{Int}_a$ *iff* $a \in \mathcal{I}(w)$
- $M, w \models \neg\varphi$ *iff not* $M, w \models \varphi$
- $M, w \models \varphi \wedge \psi$ *iff* $M, w \models \varphi$ *and* $M, w \models \psi$
- $M, w \models \mathsf{K}\varphi$ *iff* $M, v \models \varphi$ *for all* v *such that* $w \sim v$
- $M, w \models [\alpha]\psi$ *iff* $M^{\alpha}, w \models \psi$

where model M^{α} is defined according to Definitions 3 and 5 below.

Definition 3 (Update via physical action). *Given a* DL-GA *model $M = \langle W, \sim, \kappa_{\mathsf{exc}}, \mathcal{D}, \mathcal{I}, \mathcal{V} \rangle$, The update of M by a is defined as $M^a = \langle W, \sim, \kappa_{\mathsf{exc}}, \mathcal{D}, \mathcal{I}^a, \mathcal{V}^a \rangle$ where for all $w \in W$:*

$$\begin{aligned} \mathcal{I}^a(w) &= \mathcal{I}(w) \setminus \{a\} \\ \mathcal{V}^a(w) &= (\mathcal{V}(w) \cup \{p : M, w \models \gamma^+(a, p)\}) \setminus \\ &\quad\ \{p : M, w \models \gamma^-(a, p)\} \end{aligned}$$

The performance of a physical action a makes the commitment function $\mathcal{I}$ to remove a from the set of intentions. That is, if an agent intends to perform the physical action a, then after the performance of a the agent does not intend to perform a anymore. Of course, the agent may adopt intention a again by, for example, performing an intention update operation (see Section 4.1). Physical actions modify the physical facts via the positive effect preconditions and the negative effect preconditions, defined in Section 2.2. In particular, if the positive effect preconditions of action a with respect to p holds, then p will be true after the occurrence of a; if the negative effect preconditions of action a with respect to p holds, then p will be false after the occurrence of a.[2]

A sensing action updates the agent's information state by modifying the exceptionality degree of the worlds that the agent can imagine. Before defining such a model update, we follow [21] and lift the exceptionality of a possible world to the exceptionality of a formula viewed as a set of worlds.

Definition 4 (Exceptionality degree of a formula). *Let $||\varphi||_w = \{v \in W : M, v \models \varphi$ and $w \sim v\}$. The exceptionality degree of a formula φ at world w, noted $\kappa^w_{\mathsf{exc}}(\varphi)$, is defined as follows:*

$$\kappa^w_{\mathsf{exc}}(\varphi) = \begin{cases} \min_{v \in ||\varphi||_w} \kappa_{\mathsf{exc}}(v) & \text{if } ||\varphi||_w \neq \emptyset \\ \mathsf{max} & \text{if } ||\varphi||_w = \emptyset \end{cases}$$

As expected, the *plausibility* degree of a formula φ, noted $\kappa^w_{\mathsf{plaus}}(\varphi)$, is defined as $\mathsf{max} - \kappa^w_{\mathsf{exc}}(\varphi)$.

Definition 5 (Update via sensing action). *Given a* DL-GA *model $M = \langle W, \sim, \kappa_{\mathsf{exc}}, \mathcal{D}, \mathcal{I}, \mathcal{V} \rangle$. The update of M by the sensing action $*\varphi$ is defined as $M^{*\varphi} = \langle W, \sim, \kappa^{*\varphi}_{\mathsf{exc}}, \mathcal{D}, \mathcal{I}, \mathcal{V} \rangle$ such that for all w:*

$$\kappa^{*\varphi}_{\mathsf{exc}}(w) = \begin{cases} \kappa_{\mathsf{exc}}(w) - \kappa^w_{\mathsf{exc}}(\varphi) & \text{if } M, w \models \varphi \\ Cut_{\mathsf{B}}(\kappa_{\mathsf{exc}}(w) + \delta) & \text{if } M, w \models \neg\varphi \wedge \widehat{\mathsf{K}}\varphi \\ \kappa_{\mathsf{exc}}(w) & \text{if } M, w \models \mathsf{K}\neg\varphi \end{cases}$$

where $\delta \in Num \setminus \{0\}$ and

$$Cut_{\mathsf{B}}(x) = \begin{cases} x & \text{if } 0 \leq x \leq \mathsf{max} \\ \mathsf{max} & \text{if } x > \mathsf{max} \\ 0 & \text{if } x < 0 \end{cases}$$

The action of sensing that φ is true modifies the agent's beliefs as follows.

1. For every world w in which φ is true, the degree of exceptionality of w decreases from $\kappa_{\mathsf{exc}}(w)$ to $\kappa_{\mathsf{exc}}(w) - \kappa^w_{\mathsf{exc}}(\varphi)$, which is the same thing as saying that, degree of plausibility of w increases from $\mathsf{max} - \kappa_{\mathsf{exc}}(w)$ to $\mathsf{max} - (\kappa_{\mathsf{exc}}(w) - \kappa^w_{\mathsf{exc}}(\varphi))$.
2. For every world w in which φ is false:
   (a) if at w the agent can imagine a world in which φ is true, *i.e.* $\widehat{\mathsf{K}}\varphi$, then the degree of exceptionality of w increases from $\kappa_{\mathsf{exc}}(w)$ to $Cut_{\mathsf{max}}(\kappa_{\mathsf{exc}}(w) + \delta)$, which is the same thing as saying that, the degree of plausibility of w decreases from $\mathsf{max} - \kappa_{\mathsf{exc}}(w)$ to $\mathsf{max} - Cut_{\mathsf{max}}(\kappa_{\mathsf{exc}}(w) + \delta)$;

[2] Note that the order of the set theoretic operations in the definition seems to privilege negative effect preconditions; however, due to the *coherence* assumption COH_{γ} made in Section 2.2 the effects of a physical action will never be inconsistent.

(b) if at w the agent cannot imagine a world in which φ is true, *i.e.* $\mathsf{K}\neg\varphi$, then the degree of exceptionality of w does not change.

The condition 2(b) ensures that the agent's plausibility ordering over worlds does not change, if the agent learns something that he cannot imagine.[3] Cut_{B} is a minor technical device, taken from [3], which ensures that the new plausibility assignment fits into the finite set of natural numbers *Num*. The parameter δ is a *conservativeness* index which captures the agent's disposition (or personality trait) to radically change its beliefs in the light of a new evidence. More precisely, the higher is the index δ, and the higher is the agent's disposition to decrease the plausibility degree of those worlds in which the learnt fact φ is false. (When $\delta = \mathsf{max}$, the agent is minimally conservative). We assume that δ is different from 0 in order to ensure that, after learning that p is true, the agent will believe p for every proposition $p \in Atm$ (see validity (5) in Section 2.5 below).

In the sequel we write $\models_{\mathsf{DL\text{-}GA}} \varphi$ to mean that φ is *valid* in DL-GA (φ is true in all DL-GA models).

2.4 Definition of graded belief

Following [21], we define the concept of belief as a formula which is true in all worlds that are maximally plausible (or minimally exceptional).

Definition 6 (Belief, $\mathsf{B}\varphi$). *In model M at world w the agent believes that φ is true, i.e., $M, w \models \mathsf{B}\varphi$, if and only if, for every v such that $w \sim v$, if $\kappa_{\mathsf{exc}}(v) = 0$ then $M, v \models \varphi$.*

The following concept of graded belief is taken from [10]: the strength of the belief that φ is equal to the exceptionality degree of $\neg\varphi$.

Definition 7 (Graded belief, $\mathsf{B}^{\geq \mathsf{h}}\varphi$). *For all $\mathsf{h} \geq 1$, in model M at world w the agent believes that φ with strength at least h, i.e. $M, w \models \mathsf{B}^{\geq \mathsf{h}}\varphi$, if and only if, $\kappa^{w}_{\mathsf{exc}}(\neg\varphi) \geq \mathsf{h}$.*

An agent has the strong belief that φ if and only if, it believes that φ is true with maximal strength max.

Definition 8 (Strong belief, $\mathsf{SB}\varphi$). *In model M at world w the agent strongly believes that φ (or at w the agent is certain that φ is true), i.e., $M, w \models \mathsf{SB}\varphi$, if and only if $\kappa^{w}_{\mathsf{exc}}(\neg\varphi) = \mathsf{max}$.*

As the following proposition highlights, the concepts of belief, graded belief and strong belief semantically defined in Definitions 6-8 are all syntactically expressible in the logic DL-GA.

Proposition 1. *For every DL-GA model M, world w and $\mathsf{h} \in Num$ such that $\mathsf{h} \geq 1$:*

1. $M, w \models \mathsf{B}\varphi$ *iff* $M, w \models \mathsf{K}(\mathsf{exc}_0 \to \varphi)$
2. $M, w \models \mathsf{B}^{\geq \mathsf{h}}\varphi$ *iff* $M, w \models \mathsf{K}(\mathsf{exc}_{\leq \mathsf{h}-1} \to \varphi)$
3. $M, w \models \mathsf{SB}\varphi$ *iff* $M, w \models \mathsf{K}(\mathsf{exc}_{\leq \mathsf{max}-1} \to \varphi)$

where $\mathsf{exc}_{\leq \mathsf{k}} \stackrel{\mathsf{def}}{=} \bigvee_{0 \leq \mathsf{l} \leq \mathsf{k}} \mathsf{exc}_{\mathsf{l}}$ *for all* $\mathsf{k} \in Num$.

We define the dual operators in the usual way: $\widehat{\mathsf{B}}\varphi \stackrel{\mathsf{def}}{=} \neg\mathsf{B}\neg\varphi$, $\widehat{\mathsf{B}}^{\geq \mathsf{h}}\varphi \stackrel{\mathsf{def}}{=} \neg\mathsf{B}^{\geq \mathsf{h}}\neg\varphi$ and $\widehat{\mathsf{SB}}\varphi \stackrel{\mathsf{def}}{=} \neg\mathsf{SB}\neg\varphi$.

We assume that "the agent believes that φ *exactly* with strength h", *i.e.* $\mathsf{B}^{\mathsf{h}}\varphi$, if and only if the agent believes that φ with strength at least h and it is not the case that the agent believes that φ with strength at least $\mathsf{h}+1$. That is, we define:

$$\mathsf{B}^{\mathsf{h}}\varphi \stackrel{\mathsf{def}}{=} \mathsf{B}^{\geq \mathsf{h}}\varphi \wedge \neg\mathsf{B}^{\geq \mathsf{h}+1}\varphi \text{ if } 1 \leq \mathsf{h} < \mathsf{max}, \text{ and}$$
$$\mathsf{B}^{\mathsf{max}}\varphi \stackrel{\mathsf{def}}{=} \mathsf{B}^{\geq \mathsf{max}}\varphi$$

2.5 Some properties of epistemic attitudes

The following validities highlight some interesting properties of beliefs. For every $\mathsf{h}, \mathsf{k} \in Num$ such that $\mathsf{h} \geq 1$ and $\mathsf{k} \geq 1$ we have:

$$\models_{\mathsf{DL\text{-}GA}} \mathsf{K}\varphi \to \mathsf{B}^{\geq \mathsf{h}}\varphi \quad (1)$$
$$\models_{\mathsf{DL\text{-}GA}} \mathsf{B}\varphi \leftrightarrow \mathsf{B}^{\geq 1}\varphi \quad (2)$$
$$\models_{\mathsf{DL\text{-}GA}} \mathsf{SB}\varphi \leftrightarrow \mathsf{B}^{\geq \mathsf{max}}\varphi \quad (3)$$
$$\models_{\mathsf{DL\text{-}GA}} \neg(\mathsf{B}\varphi \wedge \mathsf{B}\neg\varphi) \quad (4)$$
$$\models_{\mathsf{DL\text{-}GA}} \widehat{\mathsf{K}}\varphi \to [*\varphi]\mathsf{B}\varphi \text{ if } \varphi \in Prop \quad (5)$$
$$\models_{\mathsf{DL\text{-}GA}} (\mathsf{B}^{\geq \mathsf{h}}\varphi \wedge \mathsf{B}^{\geq \mathsf{k}}\psi) \to \mathsf{B}^{\geq \min\{\mathsf{h},\mathsf{k}\}}(\varphi \wedge \psi) \quad (6)$$
$$\models_{\mathsf{DL\text{-}GA}} (\mathsf{B}^{\geq \mathsf{h}}\varphi \wedge \mathsf{B}^{\geq \mathsf{k}}\psi) \to \mathsf{B}^{\geq \max\{\mathsf{h},\mathsf{k}\}}(\varphi \vee \psi) \quad (7)$$

According to the validity (1), knowing that φ implies believing that φ with strength at least h. According to the validity (2), belief is graded belief with strength at least 1. According to the validity (3), the agent has the strong belief that φ if and only if, it believes that φ with maximal strength max. According to the validity (4) (which follows from the normality constraint $NORM_{\kappa_{\mathsf{exc}}}$ in Section 2.3), an agent cannot have inconsistent beliefs. The validity (5) highlights a basic property of belief revision in the sense of AGM theory [2]: if φ is an objective fact and the agent can imagine a world in which φ is true then, after learning that φ is true, the agent believes that φ.[4] According to the validities (6) and (7), if the agent believes that φ with strength at least h and believes that ψ with strength at least k, then the strength of the belief that $\varphi \wedge \psi$ is at least $\min\{\mathsf{h}, \mathsf{k}\}$; if the agent believes that φ with strength at least h and believes that ψ with strength at least k, then it believes $\varphi \vee \psi$ with strength at least $\max\{\mathsf{h}, \mathsf{k}\}$. Similar properties for graded belief are given in possibility theory [6].

3. EMOTIONS AND THEIR INTENSITY

We use the modal operators of graded belief and graded goal of the logic DL-GA to provide a logical analysis of emotions such as hope, fear, joy and distress with their intensities.

According to some psychological models [17, 11, 16] and computational models [9, 4] of emotions, the intensity of hope with respect to a given event is a monotonically increasing function of the degree to which the event is desirable and the likelihood of the event. That is, the higher is the desirability of the event, and the higher is the intensity of the agent's hope that this event will occur; the higher is the likelihood of the event, and the higher is the intensity of the agent's hope that this event will occur. Analogously, the intensity of fear with respect to a given event is a monotonically increasing function of the degree to which the event is undesirable and the likelihood of the event. There are several possible merging functions which satisfy these properties. For example, we could define the merging function *merge* as an average function, according to which the intensity of hope about a certain event is the average of the strength of the belief that the event will occur and the strength of the goal that it will occur. That is, for every $\mathsf{h}, \mathsf{k} \in Num$ representing respectively the strength of the belief and the strength of

[3]Note that the tree conditions 1, 2(a) and 2(b) cover all cases. Indeed, the third condition $\mathsf{K}\neg\varphi$ is equivalent to $\neg\varphi \wedge \mathsf{K}\neg\varphi$, because $\mathsf{K}\neg\varphi \to \neg\varphi$ is valid.

[4]The only difference with AGM theory is the condition $\widehat{\mathsf{K}}\varphi$. AGM assumes that new information φ must incorporated in the belief base (the so-called success postulate), whereas we here assume that φ must incorporated in the belief base *only if* the agent can imagine a world in which φ is true.

the goal, we could define $merge(\mathsf{h},\mathsf{k})$ as $\frac{\mathsf{h}+\mathsf{k}}{2}$. Another possibility is to define *merge* as a product function (also used in [9, 17]), according to which the intensity of hope about a certain event is the product of the strength of the belief that the event will occur and the strength of the goal that it will occur. Here we do not choose a specific merging function, as this would much depend on the domain of application in which the formal model has to be used.

The emotion intensity scale is defined by the following set:

$$EmoInt = \{y : \text{ there are } x_1, x_2 \in Num \text{ such that } merge(x_1,x_2) = y\}$$

As *Num* is finite, *EmoInt* is finite too.

Let us define the notions of hope and fear with their corresponding intensities. We say that the agent hopes with intensity i that its current intention to perform the action a will lead to the desirable consequence l if and only if, there are $\mathsf{h}, \mathsf{k} \in Num \setminus \{0\}$ such that $merge(\mathsf{h},\mathsf{k}) = \mathsf{i}$ and $\mathsf{h} < \mathsf{max}$ and: (1) the agent believes with strength h that the physical action a is executable and l will be true afterwards, (2) the agent wishes to achieve l with strength k, (3) the agent intends to perform the physical action a. Formally:

$$\mathsf{Hope}^{\mathsf{i}}(a,l) \stackrel{\mathsf{def}}{=} \bigvee_{\mathsf{h},\mathsf{k} \in Num\setminus\{0\}:\mathsf{h}<\mathsf{max} \text{ and } merge(\mathsf{h},\mathsf{k})=\mathsf{i}} (\mathsf{B}^{\mathsf{h}}\langle\langle a\rangle\rangle l \wedge \mathsf{AchG}^{\mathsf{k}} l \wedge \mathsf{Int}_a)$$

We say that the agent fears with intensity i that its current intention to perform the action a will lead to the undesirable consequence l if and only if, there are $\mathsf{h}, \mathsf{k} \in Num \setminus \{0\}$ such that $merge(\mathsf{h},\mathsf{k}) = \mathsf{i}$ and $\mathsf{h} < \mathsf{max}$ and: (1) the agent believes with strength h that the action a is executable and l will be true afterwards, (2) the agent wishes to avoid l with strength k, (3) the agent intends to perform the action a. Formally:

$$\mathsf{Fear}^{\mathsf{i}}(a,l) \stackrel{\mathsf{def}}{=} \bigvee_{\mathsf{h},\mathsf{k} \in Num\setminus\{0\}:\mathsf{h}<\mathsf{max} \text{ and } merge(\mathsf{h},\mathsf{k})=\mathsf{i}} (\mathsf{B}^{\mathsf{h}}\langle\langle a\rangle\rangle l \wedge \mathsf{AvdG}^{\mathsf{k}} l \wedge \mathsf{Int}_a)$$

In the preceding definitions of hope and fear, the strength of the belief is supposed to be less than max in order to distinguish hope and fear, which imply some form of uncertainty, from happiness and distress which are based on certainty. Indeed, we have that:

$$\models_{\mathsf{DL\text{-}GA}} \mathsf{Hope}^{\mathsf{i}}(a,l) \rightarrow \neg\mathsf{SB}\langle\langle a\rangle\rangle l \quad (8)$$

$$\models_{\mathsf{DL\text{-}GA}} \mathsf{Fear}^{\mathsf{i}}(a,l) \rightarrow \neg\mathsf{SB}\langle\langle a\rangle\rangle l \quad (9)$$

This means that if an agent hopes/fears that its intention to perform the action a will lead to the desirable/undesirable result l, then it is not certain about that. For example, if our robot hopes to place a container at a given target position by its transport plan, then the robot is not certain that the container will be at the target position after performing the transport plan. On the contrary, to be joyful/distressed that its current intention to perform the action a will lead to the desirable/undesirable consequence l, the agent should be *certain* that its intention to perform the action a will lead to the desirable/undesirable consequence l. This is consistent with OCC psychological model of emotions [16] according to which, while joy and distress are triggered by *actual consequences*, hope and fear are triggered by *prospective consequences* (or *prospects*). Like [9], we here interpret the term 'prospect' as synonymous of 'uncertain consequence' (in contrast with 'actual consequence' as synonymous of 'certain consequence'). The following are our definitions of joy and distress about actions:

$$\mathsf{Joy}^{\mathsf{i}}(a,l) \stackrel{\mathsf{def}}{=} \bigvee_{\mathsf{k} \in Num\setminus\{0\}: merge(\mathsf{max},\mathsf{k})=\mathsf{i}} (\mathsf{SB}\langle\langle a\rangle\rangle l \wedge \mathsf{AchG}^{\mathsf{k}} l \wedge \mathsf{Int}_a)$$

$$\mathsf{Distress}^{\mathsf{i}}(a,l) \stackrel{\mathsf{def}}{=} \bigvee_{\mathsf{k} \in Num\setminus\{0\}: merge(\mathsf{max},\mathsf{k})=\mathsf{i}} (\mathsf{SB}\langle\langle a\rangle\rangle l \wedge \mathsf{AvdG}^{\mathsf{k}} l \wedge \mathsf{Int}_a)$$

where $\mathsf{Joy}^{\mathsf{i}}(a,l)$ and $\mathsf{Distress}^{\mathsf{i}}(a,l)$ respectively mean that "the agent is joyful that its current intention to perform the action a will lead to the desirable consequence l" and "the agent is distressed that its current intention to perform the action a will lead to the undesirable consequence l". Note that, when computing the intensity of joy and distress, the belief parameter in the merging function *merge* is set to max because strong belief is equivalent to graded belief with maximal strength (validity (3) in Section 2.5).

We here distinguish distress from sadness by adding a condition to the definition of distress: the appraisal variable called *controllability* or *control potential* [19]. That is, to be sad that its current intention to perform the action a will lead to the undesirable result l, the agent should be certain that it has no control over the undesirable result l, in the sense that the agent cannot prevent l to be true — which is the same thing as saying that l will be true after every executable action of the agent —.

$$\mathsf{Sadness}^{\mathsf{i}}(a,l) \stackrel{\mathsf{def}}{=} \mathsf{Distress}^{\mathsf{i}}(a,l) \wedge \mathsf{SB}\neg\mathsf{Control}\ l$$

with

$$\mathsf{Control}\ \varphi \stackrel{\mathsf{def}}{=} \bigvee_{b \in PAct} \langle\langle b\rangle\rangle\neg\varphi$$

where $\mathsf{Control}\ \varphi$ means "the agent has control over φ" (or "the agent can prevent φ to be true"). Our definition is consistent with some psychological theories [19, 11] according to which, undesirable states of affairs that not be controlled makes one to be sad.

Example 1. *Consider again our robot which can decide to transport either container number 1 or container number 2 to a given target position. The former task is more demanding than the latter task, as container number 1 is much heavier than container number 2. In particular, the former task requires at least a full battery charge, whereas the latter requires at least a half battery charge. This means that the action of transporting container number 1 ($transport_1$) and the action of transporting container number 2 ($transport_2$) have the following positive effect preconditions with respect to the objective of placing a container at the target position (pos):*
$\gamma^+(transport_1, pos) = fullCharge,$
$\gamma^+(transport_2, pos) = fullCharge \vee halfCharge.$
Let us assume that the two actions are always executable:
$Pre(tranport_1) = Pre(tranport_2) = \top.$

Suppose that at the state w the robot intends to transport container number 1 to the target position and considers undesirable with degree k not to have any container at the target position, i.e.*,*

$$M, w \models \mathsf{AvdG}^{\mathsf{k}}\neg pos \wedge \mathsf{Int}_{transport_1}$$

Moreover, suppose that the robot is certain that in the current situation there is no container at the target position, i.e.*,*

$$M, w \models \mathsf{SB}\neg pos$$

Finally, suppose that the robot is minimally conservative in revising its beliefs, that is, $\delta = \mathsf{max}$.

The robot observes its battery load and realizes that it does not have a full battery charge but only a half battery charge. After the observation the robot will strongly believe that, if it follows its intention, it will not place any container at the target position, i.e.*,*

$$M^{*halfCharge \wedge \neg fullCharge}, w \models \mathsf{SB}\langle\langle transport_1\rangle\rangle\neg pos$$

It should be noted that the new model $M^{*halfCharge \wedge \neg fullCharge}$ is exactly the same as M except for the plausibility value κ_{exc}. This implies that we have,

$$M^{*halfCharge \wedge \neg fullCharge}, w \models \mathsf{AvdG}^{\mathsf{k}}\neg pos \wedge \mathsf{Int}_{transport_1} \wedge \mathsf{SB}\langle\langle transport_1\rangle\rangle\neg pos$$

and therefore for $merge(\mathsf{max},\mathsf{k}) = \mathsf{i}$ we have

$$M^{*halfCharge \wedge \neg fullCharge}, w \models \mathsf{Distress}^{\mathsf{i}}(transport_1, \neg pos)$$

This means that, after having observed that it only has a half battery charge, the robot is distressed with intensity i that, if it follows its current intention, then it will not succeed in placing a container at the target position.

4. FROM APPRAISAL TO COPING

In the previous section, we have characterized emotions in terms of beliefs, (achievement and avoidance) goals, and intentions, and formalized their intensities in terms of belief strength and goal strength. Emotions with high intensity influence the agent's behavior in order to cope with relevant and significant events. In general, coping can be seen as a cognitive mechanism whose aim is to discharge a certain emotion by modifying one or more of the mental attitudes (*e.g.*, beliefs, goals, intentions) that triggered the emotion [11]. For example, our robot can cope with its distress that if it follows its current intention then it will not succeed in placing a container at the target position, by reconsidering its current intention. The coping mechanism determines various types of responses, also called coping strategies. We here consider three types of coping strategies: coping strategies affecting intentions, coping strategies affecting beliefs and coping strategies affecting goals. More precisely, we consider coping strategies which deal with emotion either by forming or revising an intention to act in the world, or by changing the agent's interpretation of the situation (by changing belief strength or goal strength).

4.1 Coping strategies: syntax and semantics

We extend the logic DL-GA with three different kinds of coping strategies: (1) coping strategies affecting beliefs of the form $\varphi\uparrow^{\mathsf{B}}$ and $\varphi\downarrow^{\mathsf{B}}$, (2) coping strategies affecting goals of the form $l\uparrow^{\mathsf{D}}$ and $l\downarrow^{\mathsf{D}}$, and (3) coping strategies affecting intentions of the form $-a$ and $+a$. We call DL-GA$^+$ the resulting logic. $\varphi\uparrow^{\mathsf{B}}$ consists in increasing the strength of the belief that φ is true, while $\varphi\downarrow^{\mathsf{B}}$ consists in reducing the strength of the belief that φ is true. $l\uparrow^{\mathsf{D}}$ consists in increasing the desirability of l, while $l\downarrow^{\mathsf{D}}$ consists in reducing the desirability of l. Finally, $-a$ consists in removing the intention Int_a, while $+a$ consists in generating the intention Int_a.

The set of coping strategies is defined by the following grammar:

$$CStr \quad : \quad \beta \quad ::= \quad \varphi\uparrow^{\mathsf{B}} \mid \varphi\downarrow^{\mathsf{B}} \mid l\uparrow^{\mathsf{D}} \mid l\downarrow^{\mathsf{D}} \mid -a \mid +a$$

where φ ranges over *Fml*, l ranges over *Lit*, and a ranges over *PAct*. For every coping strategy β we introduce a corresponding dynamic operator $[\beta]$, where $[\beta]\psi$ has to be read "after the occurrence of β, ψ will be true".

As expected, the truth conditions of the new operators are given in terms of model transformation. For every $\beta \in CStr$ we define:

$$M, w \models [\beta]\psi \text{ iff } M^{\beta}, w \models \psi$$

The model M^{β} is defined according to the Definitions 9-11 below.

Coping strategies affecting the strength of the belief that φ either increase or decrease the exceptionality of the worlds in which φ is false with ω unit, only if the agent believes that φ, *i.e.* $\mathsf{B}\varphi$. If the agent does not believe that φ, *i.e.* $\neg\mathsf{B}\varphi$, they do not have any effect on the agent's mental state. In fact, we assume that coping strategies can only operate on *existing* beliefs of the agent by either increasing or decreasing their strengths. ω is a parameter which captures the agent's disposition (or personality trait) to radically change its mental state when coping with emotions (the higher is ω, and the higher is the agent's disposition to change its mental state when coping with emotions).

Definition 9 (Update via coping strategy on beliefs). *Given a DL-GA model M and $\beta \in \{\varphi\uparrow^{\mathsf{B}}, \varphi\downarrow^{\mathsf{B}}\}$, the updated model M^{β} is defined as $M^{\beta} = \langle W, \sim, \kappa^{\beta}_{\mathsf{exc}}, \mathcal{D}, \mathcal{I}, \mathcal{V}\rangle$ where for all w:*

$$\kappa^{\beta}_{\mathsf{exc}}(w) = \begin{cases} \kappa_{\mathsf{exc}}(w) & \text{if } M, w \models \varphi \\ Cut_{\mathsf{B}}(\kappa_{\mathsf{exc}}(w) + \omega) & \text{if } M, w \models \neg\varphi \wedge \mathsf{B}\varphi \text{ and } \beta = \varphi\uparrow^{\mathsf{B}} \\ \kappa_{\mathsf{exc}}(w) & \text{if } M, w \models \neg\varphi \wedge \neg\mathsf{B}\varphi \text{ and } \beta = \varphi\uparrow^{\mathsf{B}} \\ Cut_{\mathsf{B}}(\kappa_{\mathsf{exc}}(w) - \omega) & \text{if } M, w \models \neg\varphi \wedge \mathsf{B}\varphi \text{ and } \beta = \varphi\downarrow^{\mathsf{B}} \\ \kappa_{\mathsf{exc}}(w) & \text{if } M, w \models \neg\varphi \wedge \neg\mathsf{B}\varphi \text{ and } \beta = \varphi\downarrow^{\mathsf{B}} \end{cases}$$

$\omega \in Num \setminus \{0\}$ and Cut_{B} has been defined in Definition 5.

Coping strategies affecting desirability of l either increase or decrease the desirability of l with ω unit.

Definition 10 (Update via coping strategy on goals). *Given a DL-GA model M and $\beta \in \{l\uparrow^{\mathsf{D}}, l\downarrow^{\mathsf{D}}\}$, the updated model M^{β} is defined as $M^{\beta} = \langle W, \sim, \kappa_{\mathsf{exc}}, \mathcal{D}^{\beta}, \mathcal{I}, \mathcal{V}\rangle$ where for all w:*

$$\mathcal{D}^{\beta}(w, l') = \begin{cases} Cut_{\mathsf{D}}(\mathcal{D}(w, l') + \omega) & \text{if } \beta = l\uparrow^{\mathsf{D}} \text{ and } l' = l \\ Cut_{\mathsf{D}}(\mathcal{D}(w, l') - \omega) & \text{if } \beta = l\downarrow^{\mathsf{D}} \text{ and } l' = l \\ \mathcal{D}(w, l') & \text{if } l' \neq l \end{cases}$$

$\omega \in Num \setminus \{0\}$ and:

$$Cut_{\mathsf{D}}(y) = \begin{cases} y & \text{if } -\mathsf{max} \leq y \leq \mathsf{max} \\ \mathsf{max} & \text{if } y > \mathsf{max} \\ -\mathsf{max} & \text{if } y < -\mathsf{max} \end{cases}$$

Cut_{D} ensures that the new desirability degree of a literal fits into the finite set of integers $Num \cup Num^{-}$.

Finally, coping strategies affecting intentions change the commitment function by either adding or removing an intention.

Definition 11 (Update via coping strategy on intentions). *Given a DL-GA model M and $\beta \in \{-a, +a\}$, the updated model M^{β} is defined as $M^{\beta} = \langle W, \sim, \kappa_{\mathsf{exc}}, \mathcal{D}, \mathcal{I}^{\beta}, C, \mathcal{V}\rangle$ where for all w:*

$$\mathcal{I}^{\beta}(w) = \begin{cases} \mathcal{I}(w) \setminus \{a\} & \text{if } \beta = -a \\ \mathcal{I}(w) \cup \{a\} & \text{if } \beta = +a \end{cases}$$

The following validities capture some expected properties of coping strategies affecting beliefs and goals. If $\mathsf{h} \geq 1$ then:

$$\models_{\mathsf{DL\text{-}GA}^+} \mathsf{B}^{\geq \mathsf{h}}\varphi \rightarrow [\varphi\uparrow^{\mathsf{B}}]\mathsf{B}^{\geq Cut_{\mathsf{B}}(\mathsf{h}+\omega)}\varphi \qquad (10)$$

$$\models_{\mathsf{DL\text{-}GA}^+} \mathsf{B}^{\geq \mathsf{h}}\varphi \rightarrow [\varphi\downarrow^{\mathsf{B}}]\mathsf{B}^{\geq Cut_{\mathsf{B}}(\mathsf{h}-\omega)}\varphi \text{ if } Cut_{\mathsf{B}}(\mathsf{h}-\omega) > 0 \qquad (11)$$

$$\models_{\mathsf{DL\text{-}GA}^+} \mathsf{B}^{\geq \mathsf{h}}\varphi \rightarrow [\varphi\downarrow^{\mathsf{B}}]\neg\mathsf{B}\varphi \text{ if } Cut_{\mathsf{B}}(\mathsf{h}-\omega) = 0 \qquad (12)$$

$$\models_{\mathsf{DL\text{-}GA}^+} \mathsf{Des}^{\mathsf{h}} l \rightarrow [\varphi\uparrow^{\mathsf{D}}]\mathsf{Des}^{Cut_{\mathsf{D}}(\mathsf{h}+\omega)} l \qquad (13)$$

$$\models_{\mathsf{DL\text{-}GA}^+} \mathsf{Des}^{\mathsf{h}} l \rightarrow [\varphi\downarrow^{\mathsf{D}}]\mathsf{Des}^{Cut_{\mathsf{D}}(\mathsf{h}-\omega)} l \qquad (14)$$

4.2 Triggering conditions of coping strategies

In our model coping strategies have triggering conditions which are captured by the function

$$Trg : CStr \longrightarrow Fml$$

mapping coping strategies to DL-GA-formulas. For every coping strategy β, $Trg(\beta)$ captures the conditions under which the coping strategy β is *possibly* triggered. Following current psychological and computational models of emotions [11, 16, 9], we here assume that coping strategies are triggered by the agent's positively valenced emotions (*e.g.*, hope and joy) and negatively valenced emotions (*e.g.*, fear and sadness). In what follows we only discuss coping strategies triggered by negatively valenced emotions.

We assume that an agent that is fearful or distressed because its intention a will realize the undesirable effect l will possibly reconsider its intention. Such an intention reconsideration strategy can be formulated as follows:

$$Trg(-a) = \bigvee_{l \in Lit, \mathsf{i} \in EmoInt: \mathsf{i} \geq \theta}((\mathsf{Fear}^{\mathsf{i}}(a, l) \vee \mathsf{Distress}^{\mathsf{i}}(a, l)) \wedge \mathsf{B}\ \mathsf{Control}\ l)$$

This means that the coping strategy of reconsidering the intention to perform the action a is triggered if and only if (1) the agent is either fearful or distressed with intensity at least θ that its intention to perform the action a will lead to an undesirable result, (2) the agent believes that he has control over l, in the sense that he can prevent the undesirable result l to be true by performing a different action. θ is a threshold which captures the agent's sensitivity to negative emotions (the lower is θ, and the higher is the agent's disposition to discharge a negative emotion by coping with it). Control l captures the appraisal variable called *controllability* we have discussed in Section 3.

Furthermore, we assume that an agent that is fearful or distressed because it believes that its intended action a will realize the undesirable consequence l on which it has no control (1) will decrease the strength of the belief that action a will lead to the undesirable consequence l or, (2) will increase the desirability of l. The former kind of coping strategy captures *wishful thinking* while the latter captures *mental disengagement*.

$Trg(\langle\langle a\rangle\rangle l{\downarrow}^{\mathsf{B}}) = \bigvee_{l\in Lit, \mathsf{i}\in EmoInt:\mathsf{i}\geq\theta}((\mathsf{Fear}^{\mathsf{i}}(a,l) \vee \mathsf{Distress}^{\mathsf{i}}(a,l)) \wedge \neg\mathsf{B}\ \mathsf{Control}\ l)$

$Trg(l{\uparrow}^{\mathsf{D}}) = \bigvee_{l\in Lit, \mathsf{i}\in EmoInt:\mathsf{i}\geq\theta}((\mathsf{Fear}^{\mathsf{i}}(a,l) \vee \mathsf{Distress}^{\mathsf{i}}(a,l)) \wedge \neg\mathsf{B}\ \mathsf{Control}\ l)$

Note that differently from intention-related coping, wishful thinking and mental disengagement are triggered if the agent appraises that it has no controllability of the undesirable consequence l, in the sense that it cannot prevent l to be true (on this see also [14]).

Example 2. *Let us continue the example of Section 3. We have,*

$$M^{*halfCharge\wedge\neg fullCharge}, w \models \mathsf{Distress}^{\mathsf{i}}(transport_1, \neg pos)$$

Suppose $\mathsf{i} \geq \theta$. *Given the assumption that* $Pre(tranport_2) = \top$ *and the positive effect preconditions of* $transport_2$ *with respect to pos, the robot believes that the action* $transport_2$ *will place the second container at the target position,* i.e.*,*

$$M^{*halfCharge\wedge\neg fullCharge}, w \models \mathsf{B}\langle\langle transport_2\rangle\rangle pos$$

and therefore,

$$M^{*halfCharge\wedge\neg fullCharge}, w \models \mathsf{B}\ \mathsf{Control}\ \neg pos$$

Following the specification of the triggering condition for intention-related coping, the robot can now reconsider its intention $\mathsf{Int}_{transport_1}$,

$$M^{*halfCharge\wedge\neg fullCharge}, w \models Trg(-transport_1)$$

5. AXIOMATIZATION AND DECIDABILITY

The logic DL-GA of Section 2 is axiomatized as an extension of the normal modal logic **S5** for the epistemic operator K with (1) a theory describing the constraints imposed on the agent's mental state, (2) the reduction axioms of the dynamic operators $[\alpha]$, and (3) an inference rule of replacement of equivalents.

Theory of the agent's mental state.

$$\bigvee_{\mathsf{h}\in Num} \mathsf{exc}_{\mathsf{h}}$$
$$\bigvee_{\mathsf{k}\in Num\cup Num^-} \mathsf{Des}^{\mathsf{k}} l$$
$$\mathsf{exc}_{\mathsf{h}} \rightarrow \neg\mathsf{exc}_{\mathsf{l}} \text{ if } \mathsf{h}\neq\mathsf{l}$$
$$\mathsf{Des}^{\mathsf{k}} l \rightarrow \neg\mathsf{Des}^{\mathsf{m}} l \text{ if } \mathsf{k}\neq\mathsf{m}$$
$$\widehat{\mathsf{K}}\ \mathsf{exc}_0$$

Reduction axioms for the dynamic operators $[\alpha]$.

$$[\alpha]p \leftrightarrow \begin{cases} (\gamma^+(a,p) \wedge \neg\gamma^-(a,p)) \vee (p \wedge \neg\gamma^-(a,p)) & \text{if } \alpha = a \\ p \ \text{ if } \alpha = *\varphi \end{cases}$$

$$[\alpha]\mathsf{Int}_a \leftrightarrow \begin{cases} \bot \ \text{ if } \alpha = a \\ \mathsf{Int}_a \ \text{ if } \alpha \neq a \end{cases}$$

$$[\alpha]\mathsf{exc}_{\mathsf{h}} \leftrightarrow \begin{cases} \mathsf{exc}_{\mathsf{h}} \ \text{ if } \alpha = a \\ ((\varphi \wedge \bigvee_{\mathsf{l},\mathsf{m}\in Num\setminus\{0\}:\mathsf{l}-\mathsf{m}=\mathsf{h}}(\mathsf{B}^{\mathsf{m}}\neg\varphi \wedge \mathsf{exc}_{\mathsf{l}}))\vee \\ (\varphi \wedge (\widehat{\mathsf{B}}\varphi \wedge \mathsf{exc}_{\mathsf{h}}))\vee \\ (\neg\varphi \wedge \widehat{\mathsf{K}}\varphi \wedge \bigvee_{\mathsf{l}\in Num:Cut_{\mathsf{B}}(\mathsf{l}+\delta)=\mathsf{h}} \mathsf{exc}_{\mathsf{l}})\vee \\ (\mathsf{K}\neg\varphi \wedge \mathsf{exc}_{\mathsf{h}})) \text{ if } \alpha = *\varphi \end{cases}$$

$$[\alpha]\mathsf{Des}^{\mathsf{k}} l \leftrightarrow \mathsf{Des}^{\mathsf{k}} l$$
$$[\alpha]\neg\psi \leftrightarrow \neg[\alpha]\psi$$
$$[\alpha](\psi_1 \wedge \psi_2) \leftrightarrow ([\alpha]\psi_1 \wedge [\alpha]\psi_2)$$
$$[\alpha]\mathsf{K}\psi \leftrightarrow \mathsf{K}[\alpha]\psi$$

Rule of replacement of equivalents.

$$\text{From } \psi_1 \leftrightarrow \psi_2 \text{ infer } \varphi \leftrightarrow \varphi[\psi_1/\psi_2]$$

Given a formula φ, let $red(\varphi)$ be the formula obtained by iterating the application of the reduction axioms from the left to the right, starting from one of the innermost dynamic operators $[\alpha]$. *red* pushes the dynamic operators inside the formula, and finally eliminates them when facing an atomic proposition. Obviously, $red(\varphi)$ does not contain dynamic operators $[\alpha]$. The following proposition is proved using the reduction axioms above and the rule of replacement of equivalents.

Proposition 2. *Let* φ *be a formula in the language of DL-GA. Then,* $red(\varphi) \leftrightarrow \varphi$ *is DL-GA valid.*

Theorem 1. *Satisfiability in DL-GA is decidable.*

Sketch of Proof. Let L-GA be the fragment of the logic DL-GA without dynamic operators. The problem of satisfiability in L-GA is reducible to the problem of *global* logical consequence in **S5**, where the set of global axioms Γ is the theory of the agent's mental state given above. That is, we have $\models_{\text{L-GA}} \varphi$ if and only if $\Gamma \models_{\mathbf{S5}} \varphi$. Observe that Γ is finite. It is well-known that the problem of global logical consequence in **S5** with a finite number of global axioms is reducible to the problem of satisfiability in **S5**. The problem of satisfiability checking in **S5** is decidable [7]. It follows that the problem of satisfiability checking in the logic L-GA is decidable too. Proposition 2 and the fact that L-GA is a conservative extension of DL-GA ensure that *red* provides an effective procedure for reducing a DL-GA formula φ into an equivalent L-GA formula $red(\varphi)$. As L-GA is decidable, DL-GA is decidable too.

■

The logic DL-GA$^+$ of Section 4 is axiomatized by the axioms and the rules of inference of the logic DL-GA *plus* the following reduction axioms for the dynamic operators $[\beta]$.

Reduction axioms for the dynamic operators $[\beta]$.

$$[\beta]p \leftrightarrow p$$

$$[\beta]\mathsf{Int}_a \leftrightarrow \begin{cases} \top \ \text{ if } \beta = +a \\ \bot \ \text{ if } \beta = -a \\ \mathsf{Int}_a \ \text{ if } \beta \neq +a \text{ and } \beta \neq -a \end{cases}$$

$$[\beta]\mathsf{exc}_{\mathsf{h}} \leftrightarrow \begin{cases} ((\varphi \vee (\neg\varphi \wedge \neg\mathsf{B}\varphi)) \wedge \mathsf{exc}_{\mathsf{h}})\vee \\ (\neg\varphi \wedge \mathsf{B}\varphi \wedge \bigvee_{\mathsf{l}\in Num:Cut_{\mathsf{B}}(\mathsf{l}+\omega)=\mathsf{h}} \mathsf{exc}_{\mathsf{l}}) \ \text{ if } \beta = \varphi{\uparrow}^{\mathsf{B}} \\ ((\varphi \vee (\neg\varphi \wedge \neg\mathsf{B}\varphi)) \wedge \mathsf{exc}_{\mathsf{h}})\vee \\ (\neg\varphi \wedge \mathsf{B}\varphi \wedge \bigvee_{\mathsf{l}\in Num:Cut_{\mathsf{B}}(\mathsf{l}-\omega)=\mathsf{h}} \mathsf{exc}_{\mathsf{l}}) \ \text{ if } \beta = \varphi{\downarrow}^{\mathsf{B}} \\ \mathsf{exc}_{\mathsf{h}} \ \text{ if } \beta = l{\uparrow}^{\mathsf{D}} \ \text{ or } \beta = l{\downarrow}^{\mathsf{D}} \end{cases}$$

$$[\beta]\mathsf{Des}^{\mathsf{k}} l \leftrightarrow \begin{cases} \bigvee_{\mathsf{m}\in Num\cup Num^-:Cut_{\mathsf{D}}(\mathsf{m}+\omega)=\mathsf{k}} \mathsf{Des}^{\mathsf{m}} l \ \text{ if } \beta = l{\uparrow}^{\mathsf{D}} \\ \bigvee_{\mathsf{m}\in Num\cup Num^-:Cut_{\mathsf{D}}(\mathsf{m}-\omega)=\mathsf{k}} \mathsf{Des}^{\mathsf{m}} l \ \text{ if } \beta = l{\downarrow}^{\mathsf{D}} \\ \mathsf{Des}^{\mathsf{k}} l \ \text{ if } \beta \neq l{\uparrow}^{\mathsf{D}} \ \text{ and } \beta \neq l{\downarrow}^{\mathsf{D}} \end{cases}$$

$$[\beta]\neg\psi \leftrightarrow \neg[\beta]\psi$$
$$[\beta](\psi_1 \wedge \psi_2) \leftrightarrow ([\beta]\psi_1 \wedge [\beta]\psi_2)$$
$$[\beta]\mathsf{K}\psi \leftrightarrow \mathsf{K}[\beta]\psi$$

The following Theorem 2 is proved in the same way as Theorem 1.

Theorem 2. *Satisfiability in DL-GA$^+$ is decidable.*

6. RELATED WORK

Although psychological models of emotion emphasize the role of emotion intensity and its role in the coping mechanism, most existing works on logical modeling of emotions have ignored either the intensity of emotions or the coping strategies.

Adam et al. [1] have proposed a logical formalization of the OCC model, while Lorini & Schwarzentruber [13] have formalized counterfactual emotions such as regret and disappointment. Both approaches ignore the quantitative aspect of emotions. In a previous work [12] we formalized emotion intensity by using a similar logic, but we did not consider the coping strategies.

The logical approach to emotion proposed by Steunebrink et al. [22] has both characteristics of our approach: it provides a formal model of emotions extended with their intensities and coping strategies. In this model, an intensity function is assigned to each appraised emotion to determine its intensity at each state of the model. The coping mechanism introduced in this model is inspired by Frijda's theory of action tendencies [8]. According to this theory, specific emotions give agents the tendency to perform particular actions. In the proposed model, coping strategies are developed for negative emotions and their aim is to reduce the intensity of negative emotions. However, unlike the present approach, Steunebrink et al.'s approach takes emotion intensity as a primitive without explaining how it depends on more primitive cognitive ingredients such as belief strength and goal strength. The other important difference between the present work and Steunebrink et al.'s work is that we here provide a decidable logic of emotion with a complete axiomatization, whereas Steunebrink et al. do not provide any decidability result or complete axiomatization for their logic of emotion.

In the computational model proposed by Gratch and Marsella [9, 14], the eliciting conditions of emotions are defined in terms of quantitative measures such as desirability and likelihood of events. The model is based on several thresholds that determine when emotions are elicited and how emotions are coped with. The implementation of the proposed model is called EMA and is applied to generate predictions about human emotions and their coping strategies. Since the model is quantitative and the authors do no provide any details about its underlying logic, it is hard to compare this model with other logical approaches. One can only conclude that the model proposed by Gratch and Marsella considers both emotion intensities and coping strategies, although it does not provide a logical characterization of the emotions, their intensities, or the corresponding coping strategies.

7. DISCUSSION

In this work we have provided a logical characterization of emotions enriched with intensities and coping strategies. Emotions are defined in terms of graded beliefs, graded (achievement and avoidance) goals, and intentions. The intensity of emotions, which is defined as a function of belief strength and goal strength, is used to trigger specific coping strategies. We have considered only a few coping strategies triggered by negative emotions. In future work, we intend to extend our analysis to coping strategies triggered by positive emotions. For example, hope or joy with respect to a current intention may trigger coping strategies that suspend the other intentions in order to create a focus on the intended action for which the agent is hopeful or joyful. Moreover, the emotions discussed in this paper are defined with respect to an agent's action. We would like to extend our model in order to characterize emotions in terms of events that are independent from the agent's actions and intentions. Finally, in the present work we have only modeled the so-called prospective emotions, rather than actual emotions. We believe that our model can be easily extended to characterize actual emotions such as being joyful to have already placed a container at the target position or being hopeful that the current state of the battery charge is not empty.

8. REFERENCES

[1] C. Adam, A. Herzig, and D. Longin. A logical formalization of the OCC theory of emotions. *Synthese*, 168(2):201–248, 2009.

[2] C. E. Alchourrón, P. Gärdenfors, and D. Makinson. On the logic of theory change: partial meet contraction and revision functions. *J. of Symbolic Logic*, 50(2):510–530, 1985.

[3] G. Aucher. A combined system for update logic and belief revision. In *Proc. of PRIMA 2004*, LNAI, pages 1–18. Springer-Verlag, 2005.

[4] T. Bosse and E. Zwanenburg. There's always hope: Enhancing agent believability through expectation-based emotions. In *Proc. of ACII 2009*, pages 111–118. IEEE Computer Society Press, 2009.

[5] A. R. Damasio. *Descartes' Error: Emotion, Reason, and the Human Brain*. Putnam Pub Group, 1994.

[6] D. Dubois and H. Prade. *Possibility theory: an approach to computerized processing of uncertainty*. Plenum Press, 1988.

[7] R. Fagin, J. Halpern, Y. Moses, and M. Vardi. *Reasoning about Knowledge*. MIT Press, 1995.

[8] N. Frijda. *The Emotions*. Cambridge University Press, 1987.

[9] J. Gratch and S. Marsella. A domain independent framework for modeling emotion. *Cognitive Systems Research*, 5(4):269–306, 2004.

[10] N. Laverny and J. Lang. From knowledge-based programs to graded belief-based programs, part II: off-line reasoning. In *Proc. of IJCAI'05*, pages 497–502, 2005.

[11] R. S. Lazarus. *Emotion and adaptation*. Oxford University Press, 1991.

[12] E. Lorini. A dynamic logic of knowledge, graded beliefs and graded goals and its application to emotion modelling. In *Proc. of LORI-III*, LNCS, pages 165–178. Springer-Verlag, 2011.

[13] E. Lorini and F. Schwarzentruber. A logic for reasoning about counterfactual emotions. *Artificial Intelligence*, 175(3-4):814–847, 2011.

[14] S. Marsella and J. Gratch. EMA: A process model of appraisal dynamics. *Cognitive Systems Research*, 10:70–90, 2009.

[15] W. U. Meyer, R. Reisenzein, and A. Schützwohl. Towards a process analysis of emotions: The case of surprise. *Motivation and Emotion*, 21:251–274, 1997.

[16] A. Ortony, G. L. Clore, and A. Collins. *The cognitive structure of emotions*. Cambridge University Press, 1988.

[17] R. Reisenzein. Emotions as metarepresentational states of mind: naturalizing the belief-desire theory of emotion. *Cognitive Systems Research*, 10:6–20, 2009.

[18] R. Reiter. *Knowledge in action: logical foundations for specifying and implementing dynamical systems*. MIT Press, 2001.

[19] I. J. Roseman, A. A. Antoniou, and P. E. Jose. Appraisal determinants of emotions: constructing a more accurate and comprehensive theory. *Cognition and Emotion*, 10:241–277, 1996.

[20] H. A. Simon. Motivational and emotional controls of cognition. *Psychological Review*, 74:29–39, 1967.

[21] W. Spohn. Ordinal conditional functions: a dynamic theory of epistemic states. In *Causation in decision, belief change and statistics*, pages 105–134. Kluwer, 1998.

[22] B. R. Steunebrink, M. Dastani, and J.-J. Ch. Meyer. A formal model of emotion-based action tendency for intelligent agents. In *Proc. of EPIA'09*, LNAI, pages 174–186. Springer-Verlag, 2009.