CHAPTER

The Mental States of Agents 11

Many instances of commonsense reasoning involve the mental states of agents. An *agent* is an entity that performs purposeful actions in the world in which it exists. Examples of agents are people, animals, organizations, robots, and computer programs. In order to represent and reason about agent behavior—in particular, to interpret an agent's actions and make predictions about an agent's future actions—we must model the mental states of agents. In this chapter, we address agent behavior, which requires the modeling of beliefs, goals, and plans. We discuss the modeling of emotions. We describe the epistemic functional event calculus for reasoning about knowledge and action.

11.1 BELIEFS, GOALS, AND PLANS

Agent behavior can be reactive or goal-driven.

11.1.1 REACTIVE BEHAVIOR

Consider the following example of commonsense reasoning:

> *Ryan is driving along when a squirrel leaps out in front of the car. What will Ryan do? He will slam on the brakes.*

Here we are reasoning about the *reactive behavior* of an agent, in which an agent immediately performs an action in response to the situation the agent encounters. We may represent reactive behavior in the event calculus using trigger axioms, which are of the form

$$\gamma \Rightarrow Happens(\alpha, \tau)$$

where γ is a condition, α is an event term, and τ is a timepoint term.

11.1.2 GOAL-DRIVEN BEHAVIOR

Whereas some agent behavior is reactive, much behavior is goal-driven.

> *Ryan has not eaten for a while. He gets hungry. He walks into the kitchen. The refrigerator is empty. He goes to a sandwich shop across the street. He eats a sandwich.*

Here Ryan initiates a goal to eat some food in response to his hunger. He forms a plan to eat some food in the kitchen. We use these notions of goals and plans to interpret Ryan's actions and predict his future actions. We understand that Ryan walks into the kitchen because of his goal to eat. We understand that his plan involves taking some food out of the refrigerator. He is acting on his belief that there is food in the refrigerator. When he learns the refrigerator is empty, we expect him to revise his plan and go somewhere else.

Intuitively, a *belief* of an agent is a property considered by the agent to be true or false. A *goal* of an agent is a property that the agent desires to be true. An agent's *plan* for a goal consists of actions that the agent intends to perform in order to achieve the goal. An agent activates certain goals in certain situations. An agent's goals tend to persist until they are achieved. When an agent has a goal, the agent activates a plan for achieving the goal. When a plan is active, an agent tries to execute the plan. If a plan fails, an agent activates an alternative plan. When a goal succeeds, whether intentionally or accidentally, the goal and plan are dropped.

11.1.3 FORMALIZATION

We now formalize these notions. We represent a plan as a list of actions. We introduce a list sort and terms of the form

$$[\tau_1, \ldots, \tau_n] \tag{11.1}$$

where $\tau_1, \ldots, \tau_n$ are terms. Several functions can be applied to a list: The *Head* of (11.1) is τ_1, its *Tail* is $[\tau_2, \ldots, \tau_n]$, and its *Length* is n. An axiomatization of lists is provided in Section A.6.2.

We then use the following sorts:

- an agent sort, with variables a, a_1, a_2, ...
- a belief sort, which is a subsort of the fluent sort, with variables b, b_1, b_2, ...
- a goal sort, which is a subsort of the fluent sort, with variables g, g_1, g_2, ...
- a plan sort, which is a subsort of the list sort, with variables p, p_1, p_2, ...

In the event calculus, it is natural to represent persistent goals and plans as fluents subject to the commonsense law of inertia. We represent beliefs, goals, and plans using the following fluents:

$Believe(a, b)$: Agent a has belief b; that is, a believes that b is true.
$Goal(a, g)$: Agent a has an active goal g.
$Plan(a, g, p)$: Agent a has an active plan p for an active goal g.

We may thus write, for example, $HoldsAt(Believe(Agent, At(Agent, Kitchen)), 0)$.

Beliefs, goals, and plans may be added (activated) and dropped (deactivated).

Axiom A1.

$$Initiates(AddBelief(a, b), Believe(a, b), t)$$

Axiom A2.

$$Terminates(DropBelief(a, b), Believe(a, b), t)$$

Axiom A3.

$$Initiates(AddGoal(a,g), Goal(a,g), t)$$

Axiom A4.

$$Terminates(DropGoal(a,g), Goal(a,g), t)$$

Axiom A5.

$$Initiates(AddPlan(a,g,p), Plan(a,g,p), t)$$

Axiom A6.

$$Terminates(DropPlan(a,g,p), Plan(a,g,p), t)$$

Plans are activated for goals as follows. If an agent has a goal, the agent does not believe the goal is true, and no plans are yet active for the goal, then the agent selects and activates a plan for the goal.

Axiom A7.

$$\begin{gathered} HoldsAt(Goal(a,g),t) \land \\ \neg HoldsAt(Believe(a,g),t) \land \\ SelectedPlan(a,g,p,t) \land \\ \neg \exists p_1\, HoldsAt(Plan(a,g,p_1),t) \Rightarrow \\ Happens(AddPlan(a,g,p),t) \end{gathered}$$

We use $SelectedPlan(a,g,p,t)$ to mean that at time t, plan p is the plan selected by agent a for goal g. For a particular domain or problem, we must supply a definition for $SelectedPlan(a,g,p,t)$. This predicate represents the agent's planning knowledge, which is assumed to be fixed over time. If we wish to allow an agent to learn or modify its planning knowledge over time, then we must use a fluent instead of a predicate.

Plans are executed and managed as follows. At each timepoint, provided that the agent believes the plan to be sound and the agent does not believe the goal to be true, the agent executes the next action of the plan, and the plan is updated to consist of the remaining actions. Note that Axioms A9 and A10 together update the plan: Axiom A9 drops the current list of planned actions at the same time that Axiom A10 adds a new list of planned actions that is the tail of the current list (provided that the tail is not empty and the other conditions are met).

Axiom A8.

$$\begin{gathered} HoldsAt(Plan(a,g,p),t) \land \\ \neg HoldsAt(Believe(a,g),t) \land \\ SoundPlan(a,g,p,t) \Rightarrow \\ Happens(Head(p),t) \end{gathered}$$

Axiom A9.

$$HoldsAt(Plan(a,g,p),t) \Rightarrow Happens(DropPlan(a,g,p),t)$$

Axiom A10.

$$\begin{gathered} HoldsAt(Plan(a, g, p), t) \wedge \\ Length(p) > 1 \wedge \\ \neg HoldsAt(Believe(a, g), t) \wedge \\ SoundPlan(a, g, p, t) \Rightarrow \\ Happens(AddPlan(a, g, Tail(p)), t) \end{gathered}$$

We use $SoundPlan(a, g, p, t)$ to mean that at time t, agent a believes that p is a sound plan for achieving goal g. Notice what happens if the agent believes the plan to be unsound or believes the goal to be true: The plan is dropped. If the goal is false, then a new plan for the goal is activated by Axiom A7. For a particular domain or problem, we must supply a definition for $SoundPlan(a, g, p, t)$.

This formalization does not deal with multiple active goals, which lead to multiple active plans, which lead to the concurrent execution of plan actions, which may lead to inconsistency (see Exercise 11.7).

Goal success is treated as follows. If an agent has an active goal and the agent believes that the goal is true, then the agent drops the goal.

Axiom A11.

$$\begin{gathered} HoldsAt(Goal(a, g), t) \wedge HoldsAt(Believe(a, g), t) \Rightarrow \\ Happens(DropGoal(a, g), t) \end{gathered}$$

In order to reason about an agent in a particular domain or problem using Axioms A1 through A11, we must supply the following:

- *Event calculus axioms*: These represent the domain.
- *Belief revision axioms*: These specify how an agent's beliefs are updated in response to events and are typically of the form:

  $$\gamma \Rightarrow Initiates(\alpha, Believe(a, \beta), \tau)$$

  and

  $$\gamma \Rightarrow Terminates(\alpha, Believe(a, \beta), \tau)$$

  where γ is a condition, α is an event term, β is a fluent term, a is an agent term, and τ is a timepoint term.
- *Goal activation axioms*: These are typically of the form

  $$\gamma \Rightarrow Happens(AddGoal(a, \beta), \tau)$$

  where γ is a condition, a is an agent term, β is a fluent term, and τ is a timepoint term.
- *Planning axioms*: These supply definitions for $SelectedPlan(a, g, p, t)$ and $SoundPlan(a, g, p, t)$.

11.1.4 EXAMPLE: HUNGRY CAT SCENARIO

This example is due to Michael Winikoff, Lin Padgham, James Harland, and John Thangarajah.

A cat is sitting on the floor. He is hungry and would love to eat a cookie sitting on a nearby table. He plans to jump onto a nearby chair and, from the chair, onto the table. But after the cat jumps onto the chair, someone moves the chair away from the table. The cat then modifies his plan. He plans to jump onto a nearby shelf and then onto the table. When he jumps onto the shelf, he notices a cupcake lying there, so he gobbles it up immediately and is satiated.

We start by providing the event calculus axioms for the domain. We use a simple spatial theory. If an agent is on surface s_1, it is possible to jump from s_1 to surface s_2, and the agent jumps from s_1 to s_2, then the agent will be on s_2:

$$\begin{gathered} HoldsAt(On(a,s_1),t) \wedge HoldsAt(CanJump(s_1,s_2),t) \Rightarrow \\ Initiates(Jump(a,s_1,s_2),On(a,s_2),t) \end{gathered} \tag{11.2}$$

In this case, the agent will no longer be on surface s_1:

$$\begin{gathered} HoldsAt(On(a,s_1),t) \wedge HoldsAt(CanJump(s_1,s_2),t) \Rightarrow \\ Terminates(Jump(a,s_1,s_2),On(a,s_1),t) \end{gathered} \tag{11.3}$$

If surface s_1 is moved from surface s_2 to surface s_3, then it will be possible to jump from s_1 to s_3:

$$Initiates(Move(s_1,s_2,s_3),CanJump(s_1,s_3),t) \tag{11.4}$$

In this case, it will no longer be possible to jump from surface s_1 to surface s_2:

$$Terminates(Move(s_1,s_2,s_3),CanJump(s_1,s_2),t) \tag{11.5}$$

We use a simple theory of eating. If an agent is on a surface, some food is on the surface, and the agent eats the food, then the agent will be satiated:

$$\begin{gathered} HoldsAt(On(a,s),t) \wedge HoldsAt(On(f,s),t) \Rightarrow \\ Initiates(Eat(a,f),Satiated(a),t) \end{gathered} \tag{11.6}$$

If an agent eats food that is on a surface, then the food will no longer be on the surface:

$$\begin{gathered} HoldsAt(On(a,s),t) \wedge HoldsAt(On(f,s),t) \Rightarrow \\ Terminates(Eat(a,f),On(f,s),t) \end{gathered} \tag{11.7}$$

We provide axioms for belief revision:

$$\begin{gathered} HoldsAt(Believe(a,On(a,s_1)),t) \wedge \\ HoldsAt(Believe(a,CanJump(s_1,s_2)),t) \Rightarrow \\ Initiates(Jump(a,s_1,s_2),Believe(a,On(a,s_2)),t) \end{gathered} \tag{11.8}$$

$$\begin{gathered} HoldsAt(Believe(a,On(a,s_1)),t) \wedge \\ HoldsAt(Believe(a,CanJump(s_1,s_2)),t) \Rightarrow \\ Terminates(Jump(a,s_1,s_2),Believe(a,On(a,s_1)),t) \end{gathered} \tag{11.9}$$

$$Initiates(Move(s_1,s_2,s_3),Believe(a,CanJump(s_1,s_3)),t) \tag{11.10}$$

$$Terminates(Move(s_1,s_2,s_3),Believe(a,CanJump(s_1,s_2)),t) \tag{11.11}$$

$$HoldsAt(Believe(a, On(a, s)), t) \wedge \\ HoldsAt(Believe(a, On(f, s)), t) \Rightarrow \\ Initiates(Eat(a, f), Believe(a, Satiated(a)), t) \tag{11.12}$$

$$HoldsAt(Believe(a, On(a, s)), t) \wedge \\ HoldsAt(Believe(a, On(f, s)), t) \Rightarrow \\ Terminates(Eat(a, f), Believe(a, On(f, s)), t) \tag{11.13}$$

We provide planning axioms. We give a definition for $SelectedPlan(a, g, p, t)$ that enables the agent to plan trips consisting of one and two jumps in order to obtain food:

$$\begin{aligned} & SelectedPlan(a, g, p, t) \Leftrightarrow \\ & \exists s_1, s_2, f\ (HoldsAt(Believe(a, On(a, s_1)), t) \wedge \\ & HoldsAt(Believe(a, CanJump(s_1, s_2)), t) \wedge \\ & HoldsAt(Believe(a, On(f, s_2)), t) \wedge \\ & g = Satiated(a) \wedge \\ & p = [Jump(a, s_1, s_2), Eat(a, f)]) \vee \\ & \exists s_1, s_2, s_3, f\ (HoldsAt(Believe(a, On(a, s_1)), t) \wedge \\ & HoldsAt(Believe(a, CanJump(s_1, s_2)), t) \wedge \\ & HoldsAt(Believe(a, CanJump(s_2, s_3)), t) \wedge \\ & HoldsAt(Believe(a, On(f, s_3)), t) \wedge \\ & g = Satiated(a) \wedge \\ & p = [Jump(a, s_1, s_2), Wait(a), Jump(a, s_2, s_3), Eat(a, f)]) \end{aligned} \tag{11.14}$$

We give a definition for $SoundPlan(a, g, p, t)$ that deals with the plans selected by (11.14). An agent's plan is sound if and only if for every jumping action in the plan, the agent believes it is possible to jump from the source surface to the destination surface:

$$SoundPlan(a, g, p, t) \Leftrightarrow \\ \forall s_1, s_2\ (Jump(a, s_1, s_2) \in p \Rightarrow \\ HoldsAt(Believe(a, CanJump(s_1, s_2)), t)) \tag{11.15}$$

In addition to the goal-driven behavior, the agent also has a reactive behavior. If the agent is not satiated, the agent is on a surface, and some food is on the surface, then the agent will eat the food:

$$\neg HoldsAt(Satiated(a), t) \wedge \\ HoldsAt(On(a, s), t) \wedge \\ HoldsAt(On(f, s), t) \Rightarrow \\ Happens(Eat(a, f), t) \tag{11.16}$$

We specify some observations. The cat has the goal of being satiated, and no plans are active:

$$\neg ReleasedAt(f, t) \tag{11.17}$$

$$\begin{aligned}&HoldsAt(Goal(a,g),0) \Leftrightarrow \\ &a = Cat \wedge g = Satiated(Cat)\end{aligned} \tag{11.18}$$

$$\neg HoldsAt(Plan(a,g,p),0) \tag{11.19}$$

The cat is on the floor, food is on the table and the shelf, and it is possible to jump from the floor to the chair, from the chair to the table, and from the shelf to the table:

$$\begin{aligned}&HoldsAt(On(o,s),0) \Leftrightarrow \\ &(o = Cat \wedge s = Floor) \vee \\ &(o = Food1 \wedge s = Table) \vee \\ &(o = Food2 \wedge s = Shelf)\end{aligned} \tag{11.20}$$

$$\begin{aligned}&HoldsAt(CanJump(s_1,s_2),0) \Leftrightarrow \\ &(s_1 = Floor \wedge s_2 = Chair) \vee \\ &(s_1 = Chair \wedge s_2 = Table) \vee \\ &(s_1 = Shelf \wedge s_2 = Table)\end{aligned} \tag{11.21}$$

The cat believes the cat is on the floor, believes food is on the table, does not believe food is on the shelf, and believes it is possible to jump from the floor to the chair, from the chair to the table, and from the shelf to the table:

$$\begin{aligned}&HoldsAt(Believe(a,On(o,s)),0) \Leftrightarrow \\ &(a = Cat \wedge o = Cat \wedge s = Floor) \vee \\ &(a = Cat \wedge o = Food1 \wedge s = Table)\end{aligned} \tag{11.22}$$

$$\begin{aligned}&HoldsAt(Believe(a,CanJump(s_1,s_2)),0) \Leftrightarrow \\ &(a = Cat \wedge s_1 = Floor \wedge s_2 = Chair) \vee \\ &(a = Cat \wedge s_1 = Chair \wedge s_2 = Table) \vee \\ &(a = Cat \wedge s_1 = Shelf \wedge s_2 = Table)\end{aligned} \tag{11.23}$$

The cat does not believe the cat is satiated:

$$\neg HoldsAt(Believe(Cat,Satiated(Cat)),0) \tag{11.24}$$

The narrative consists of someone moving the chair at timepoint 2:

$$Happens(Move(Chair,Table,Shelf),2) \tag{11.25}$$

We can show that the cat will eat the food on the shelf.

Proposition 11.1. *Let* Σ *be the conjunction of A1 through A6, (11.2), (11.3), (11.4), (11.5), (11.6), (11.7), (11.8), (11.9), (11.10), (11.11), (11.12), and (11.13). Let* Δ *be the conjunction of A7 through A11, (11.16), and (11.25). Let* $\Omega = U[AddBelief, DropBelief, AddGoal, DropGoal, AddPlan, DropPlan, Jump, Move, Eat] \wedge U[Believe, Goal, Plan, On, CanJump, Satiated]$. *Let* Γ *be the conjunction of L1 through L8 (the axiomatization of lists of Section A.6.2), (11.14), (11.15), (11.17), (11.18), (11.19), (11.20), (11.21), (11.22), (11.23), and (11.24). Then we have*

$$CIRC[\Sigma; Initiates, Terminates, Releases] \wedge CIRC[\Delta; Happens] \wedge$$
$$\Omega \wedge \Gamma \wedge \text{DEC} \vdash Happens(Eat(Cat, Food2), 6)$$

Proof. From $CIRC[\Sigma; Initiates, Terminates, Releases]$ and Theorems 2.2 and 2.1, we have

$$\begin{aligned}
&Initiates(e,f,t) \Leftrightarrow \qquad (11.26)\\
&\exists a, b\,(e = AddBelief(a,b) \wedge f = Believe(a,b)) \vee\\
&\exists a, g\,(e = AddGoal(a,g) \wedge f = Goal(a,g)) \vee\\
&\exists a, g, p\,(e = AddPlan(a,g,p) \wedge f = Plan(a,g,p)) \vee\\
&\exists a, s_1, s_2\,(e = Jump(a,s_1,s_2) \wedge\\
&f = On(a,s_2) \wedge\\
&HoldsAt(On(a,s_1),t) \wedge\\
&HoldsAt(CanJump(s_1,s_2),t)) \vee\\
&\exists s_1, s_2, s_3\,(e = Move(s_1,s_2,s_3) \wedge f = CanJump(s_1,s_3)) \vee\\
&\exists a, f, s\,(e = Eat(a,f) \wedge\\
&f = Satiated(a) \wedge\\
&HoldsAt(On(a,s),t) \wedge\\
&HoldsAt(On(f,s),t)) \vee\\
&\exists a, s_1, s_2\,(e = Jump(a,s_1,s_2) \wedge\\
&f = Believe(a, On(a,s_2)) \wedge\\
&HoldsAt(Believe(a, On(a,s_1)),t) \wedge\\
&HoldsAt(Believe(a, CanJump(s_1,s_2)),t)) \vee\\
&\exists a, s_1, s_2, s_3\,(e = Move(s_1,s_2,s_3) \wedge f = Believe(a, CanJump(s_1,s_3))) \vee\\
&\exists a, f\,(e = Eat(a,f) \wedge\\
&f = Believe(a, Satiated(a)) \wedge\\
&HoldsAt(Believe(a, On(a,s)),t) \wedge\\
&HoldsAt(Believe(a, On(f,s)),t))
\end{aligned}$$

$$\begin{aligned}
&Terminates(e,f,t) \Leftrightarrow \qquad (11.27)\\
&\exists a, b\,(e = DropBelief(a,b) \wedge f = Believe(a,b)) \vee\\
&\exists a, g\,(e = DropGoal(a,g) \wedge f = Goal(a,g)) \vee\\
&\exists a, g, p\,(e = DropPlan(a,g,p) \wedge f = Plan(a,g,p)) \vee\\
&\exists a, s_1, s_2\\
&(e = Jump(a,s_1,s_2) \wedge\\
&f = On(a,s_1) \wedge\\
&HoldsAt(On(a,s_1),t) \wedge\\
&HoldsAt(CanJump(s_1,s_2),t)) \vee\\
&\exists s_1, s_2, s_3\,(e = Move(s_1,s_2,s_3) \wedge f = CanJump(s_1,s_2)) \vee\\
&\exists a, f, s\\
&(e = Eat(a,f) \wedge\\
&f = On(f,s) \wedge
\end{aligned}$$

$$
\begin{aligned}
&HoldsAt(On(a,s),t) \land \\
&HoldsAt(On(f,s),t)) \lor \\
&\exists a, s_1, s_2\, (e = Jump(a,s_1,s_2) \land \\
&f = Believe(a, On(a,s_1)) \land \\
&HoldsAt(Believe(a, On(a,s_1)),t) \land \\
&HoldsAt(Believe(a, CanJump(s_1,s_2)),t)) \lor \\
&\exists a, s_1, s_2, s_3\, (e = Move(s_1,s_2,s_3) \land f = Believe(a, CanJump(s_1,s_2))) \lor \\
&\exists a, f, s\, (e = Eat(a,f) \land \\
&f = Believe(a, On(f,s)) \land \\
&HoldsAt(Believe(a, On(a,s)),t) \land \\
&HoldsAt(Believe(a, On(f,s)),t))
\end{aligned}
$$

$$\neg Releases(e,f,t) \tag{11.28}$$

From $CIRC[\Delta; Happens]$ and Theorem 2.1, we have

$$
\begin{aligned}
&Happens(e,t) \Leftrightarrow \\
&\exists a, g, p\, (e = AddPlan(a,g,p) \land \\
&\neg HoldsAt(Believe(a,g),t) \land \\
&HoldsAt(Goal(a,g),t) \land \\
&SelectedPlan(a,g,p,t) \land \\
&\neg \exists p_1\, HoldsAt(Plan(a,g,p_1),t)) \lor \\
&\exists a, g, p\, (e = Head(p) \land \\
&HoldsAt(Plan(a,g,p),t) \land \\
&\neg HoldsAt(Believe(a,g),t) \land \\
&SoundPlan(a,g,p,t)) \lor \\
&\exists a, g, p\, (e = DropPlan(a,g,p) \land HoldsAt(Plan(a,g,p),t)) \lor \\
&\exists a, g, p\, (e = AddPlan(a,g,Tail(p)) \land \\
&HoldsAt(Plan(a,g,p),t) \land \\
&Length(p) > 1 \land \\
&\neg HoldsAt(Believe(a,g),t) \land \\
&SoundPlan(a,g,p,t)) \lor \\
&\exists a, g\, (e = DropGoal(a,g) \land \\
&HoldsAt(Goal(a,g),t) \land \\
&HoldsAt(Believe(a,g),t)) \lor \\
&\exists a, f, s\, (e = Eat(a,f) \land \\
&\neg HoldsAt(Satiated(a),t) \land \\
&HoldsAt(On(a,s),t) \land \\
&HoldsAt(On(f,s),t)) \lor \\
&(e = Move(Chair, Table, Shelf) \land t = 2)
\end{aligned}
\tag{11.29}
$$

We proceed one timepoint at a time, starting with timepoint 0. We start by determining what events occur at timepoint 0.

Timepoint 0: Events Let $P_1 = [Jump(Cat, Floor, Chair),\ Wait(Cat),\ Jump(Cat, Chair, Table),\ Eat(Cat, Food1)]$. Using (11.29), we proceed to show that one event occurs:

$$\begin{aligned} &Happens(e,0) \Leftrightarrow \\ &e = AddPlan(Cat, Satiated(Cat), P_1) \end{aligned} \tag{11.30}$$

From (11.19), we have

$$\begin{aligned} &\neg\exists e,a,g,p\,(e = Head(p) \wedge \\ &HoldsAt(Plan(a,g,p),0) \wedge \\ &\neg HoldsAt(Believe(a,g),0) \wedge \\ &SoundPlan(a,g,p,0)) \end{aligned} \tag{11.31}$$

$$\neg\exists e,a,g,p\,(e = DropPlan(a,g,p) \wedge HoldsAt(Plan(a,g,p),0)) \tag{11.32}$$

$$\begin{aligned} &\neg\exists e,a,g,p\,(e = AddPlan(a,g,Tail(p)) \wedge \\ &HoldsAt(Plan(a,g,p),0) \wedge \\ &Length(p) > 1 \wedge \\ &\neg HoldsAt(Believe(a,g),0) \wedge \\ &SoundPlan(a,g,p,0)) \end{aligned} \tag{11.33}$$

From (11.18) and (11.24), we have

$$\begin{aligned} &\neg\exists e,a,g\,(e = DropGoal(a,g) \wedge \\ &HoldsAt(Goal(a,g),0) \wedge \\ &HoldsAt(Believe(a,g),0)) \end{aligned} \tag{11.34}$$

From (11.20), we have

$$\begin{aligned} &\neg\exists e,a,f,s\,(e = Eat(a,f) \wedge \\ &\neg HoldsAt(Satiated(a),0) \wedge \\ &HoldsAt(On(a,s),0) \wedge \\ &HoldsAt(On(f,s),0)) \end{aligned} \tag{11.35}$$

From (11.22), (11.23), and (11.14), we have

$$SelectedPlan(a,g,p,0) \Leftrightarrow a = Cat \wedge g = Satiated(Cat) \wedge p = P_1$$

From this, (11.18), (11.19), (11.24), (11.31), (11.32), (11.33), (11.34), (11.35), and (11.29), we have (11.30).

Next we determine the values of relevant fluents at timepoint 1.

Timepoint 1: Fluents From (11.30) and (11.27), we have

$$\neg\exists e,o,s\,(Happens(e,0) \wedge Terminates(e, On(o,s),0))$$

From this, (11.20), (11.17), and DEC5, we have

$$HoldsAt(On(Cat, Floor),1) \tag{11.36}$$

$$HoldsAt(On(Food1, Table),1) \tag{11.37}$$

$$HoldsAt(On(Food2, Shelf),1) \tag{11.38}$$

From (11.30) and (11.26), we have

$$\neg\exists e, o, s\,(Happens(e, 0) \wedge Initiates(e, On(o, s), 0))$$

From this, (11.20), (11.17), and DEC6, we have

$$\begin{aligned}&\neg(o = Cat \wedge s = Floor) \wedge \\ &\neg(o = Food1 \wedge s = Table) \wedge \\ &\neg(o = Food2 \wedge s = Shelf\,) \Rightarrow \\ &\neg HoldsAt(On(o, s), 1)\end{aligned} \tag{11.39}$$

From (11.30) and (11.26), we have

$$\neg\exists e\,(Happens(e, 0) \wedge Initiates(e, Believe(Cat, Satiated(Cat)), 0))$$

From this, (11.24), (11.17), and DEC6, we have

$$\neg HoldsAt(Believe(Cat, Satiated(Cat)), 1) \tag{11.40}$$

From (11.30) and (11.27), we have

$$\neg\exists e, a, s_1, s_2\,(Happens(e, 0) \wedge Terminates(e, Believe(a, CanJump(s_1, s_2)), 0))$$

From this, (11.23), (11.17), and DEC5, we have

$$HoldsAt(Believe(Cat, CanJump(Floor, Chair)), 1) \tag{11.41}$$

$$HoldsAt(Believe(Cat, CanJump(Chair, Table)), 1) \tag{11.42}$$

$$HoldsAt(Believe(Cat, CanJump(Shelf\,, Table)), 1) \tag{11.43}$$

From (11.30), A5 (which follows from (11.26)), and DEC9, we have

$$HoldsAt(Plan(Cat, Satiated(Cat), P_1), 1) \tag{11.44}$$

We proceed to show the following:

$$\begin{aligned}&a \neq Cat \vee g \neq Satiated(Cat) \vee p \neq P_1 \Rightarrow \\ &\neg HoldsAt(Plan(a, g, p), 1)\end{aligned} \tag{11.45}$$

Let α be an arbitrary agent, γ be an arbitrary fluent, and π be an arbitrary list. We must show

$$\begin{aligned}&\alpha \neq Cat \vee \gamma \neq Satiated(Cat) \vee \pi \neq P_1 \Rightarrow \\ &\neg HoldsAt(Plan(\alpha, \gamma, \pi), 1)\end{aligned}$$

Suppose $\alpha \neq Cat \vee \gamma \neq Satiated(Cat) \vee \pi \neq P_1$. From this, (11.30), and (11.26), we have

$$\neg\exists e\,(Happens(e, 0) \wedge Initiates(e, Plan(\alpha, \gamma, \pi), 0))$$

From this, (11.17), (11.19), and DEC6, we have $\neg HoldsAt(Plan(\alpha, \gamma, \pi), 1)$ as required.

From (11.30) and (11.27), we have

$$\neg\exists e\,(Happens(e, 0) \wedge Terminates(e, Goal(Cat, Satiated(Cat)), 0))$$

From this, (11.18), (11.17), and DEC5, we have

$$HoldsAt(Goal(Cat, Satiated(Cat)), 1) \tag{11.46}$$

We proceed to show

$$a \neq Cat \lor g \neq Satiated(Cat) \Rightarrow \neg HoldsAt(Goal(a,g),1) \tag{11.47}$$

Let α be an arbitrary agent and γ be an arbitrary fluent. We must show

$$\alpha \neq Cat \lor \gamma \neq Satiated(Cat) \Rightarrow \neg HoldsAt(Goal(\alpha,\gamma),1)$$

Suppose $\alpha \neq Cat \lor \gamma \neq Satiated(Cat)$. From (11.30) and (11.26), we have

$$\neg\exists e\,(Happens(e,0) \land Initiates(e, Goal(\alpha,\gamma),0))$$

From this, (11.18), $\alpha \neq Cat \lor \gamma \neq Satiated(Cat)$, (11.17), and DEC6, we have $\neg HoldsAt(Goal(\alpha,\gamma),1)$ as required.

Timepoint 1: Events Using (11.29), we proceed to show

$$\begin{aligned} &Happens(e,1) \Leftrightarrow \\ &e = Jump(Cat, Floor, Chair) \lor \\ &e = DropPlan(Cat, Satiated(Cat), P_1) \lor \\ &e = AddPlan(Cat, Satiated(Cat), Tail(P_1)) \end{aligned} \tag{11.48}$$

From (11.46), (11.47), and (11.44), we have

$$\begin{aligned} &\neg\exists a,g,p\,(e = AddPlan(a,g,p) \land \\ &\neg HoldsAt(Believe(a,g),1) \land \\ &HoldsAt(Goal(a,g),1) \land \\ &SelectedPlan(a,g,p,1) \land \\ &\neg\exists p_1\, HoldsAt(Plan(a,g,p_1),1)) \end{aligned} \tag{11.49}$$

From (11.40) and (11.47), we have

$$\begin{aligned} &\neg\exists e,a,g\,(e = DropGoal(a,g) \land \\ &HoldsAt(Goal(a,g),1) \land \\ &HoldsAt(Believe(a,g),1)) \end{aligned} \tag{11.50}$$

From (11.36), (11.37), (11.38), and (11.39), we have

$$\begin{aligned} &\neg\exists e,a,f,s\,(e = Eat(a,f) \land \\ &\neg HoldsAt(Satiated(a),1) \land \\ &HoldsAt(On(a,s),1) \land \\ &HoldsAt(On(f,s),1)) \end{aligned} \tag{11.51}$$

From (11.41), (11.42), (11.43), (11.15), L5, and L6, we have

$$SoundPlan(Cat, Satiated(Cat), P_1, 1)$$

From this, (11.49), (11.50), (11.51), (11.45), (11.44), (11.40), (11.29), L1, L3, and L4, we have (11.48).

Timepoint 2: Fluents Using similar arguments we can show

$$\begin{aligned} &HoldsAt(On(o,s),2) \Leftrightarrow \\ &(o = Cat \land s = Chair) \lor \end{aligned} \tag{11.52}$$

$$
\begin{aligned}
&(o = \mathit{Food1} \land s = \mathit{Table}) \lor \\
&(o = \mathit{Food2} \land s = \mathit{Shelf}\,)
\end{aligned}
$$

$$
\begin{gathered}
\mathit{HoldsAt}(\mathit{Believe}(a, \mathit{CanJump}(s_1, s_2)), 1) \Leftrightarrow \\
(a = \mathit{Cat} \land s_1 = \mathit{Floor} \land s_2 = \mathit{Chair}) \lor \\
(a = \mathit{Cat} \land s_1 = \mathit{Chair} \land s_2 = \mathit{Table}) \lor \\
(a = \mathit{Cat} \land s_1 = \mathit{Shelf} \land s_2 = \mathit{Table})
\end{gathered} \tag{11.53}
$$

$$
\neg \mathit{HoldsAt}(\mathit{Believe}(\mathit{Cat}, \mathit{Satiated}(\mathit{Cat})), 2) \tag{11.54}
$$

$$
\mathit{HoldsAt}(\mathit{Plan}(\mathit{Cat}, \mathit{Satiated}(\mathit{Cat}), \mathit{Tail}(P_1)), 2) \tag{11.55}
$$

$$
\begin{gathered}
a \neq \mathit{Cat} \lor g \neq \mathit{Satiated}(\mathit{Cat}) \lor p \neq \mathit{Tail}(P_1) \Rightarrow \\
\neg \mathit{HoldsAt}(\mathit{Plan}(a, g, p), 2)
\end{gathered} \tag{11.56}
$$

$$
\mathit{HoldsAt}(\mathit{Goal}(\mathit{Cat}, \mathit{Satiated}(\mathit{Cat})), 2) \tag{11.57}
$$

$$
a \neq \mathit{Cat} \lor g \neq \mathit{Satiated}(\mathit{Cat}) \Rightarrow \neg \mathit{HoldsAt}(\mathit{Goal}(a, g), 2) \tag{11.58}
$$

Timepoint 2: Events We proceed to show

$$
\begin{gathered}
\mathit{Happens}(e, 2) \Leftrightarrow \\
e = \mathit{Move}(\mathit{Chair}, \mathit{Table}, \mathit{Shelf}\,) \lor \\
e = \mathit{Wait}(\mathit{Cat}) \lor \\
e = \mathit{DropPlan}(\mathit{Cat}, \mathit{Satiated}(\mathit{Cat}), \mathit{Tail}(P_1)) \lor \\
e = \mathit{AddPlan}(\mathit{Cat}, \mathit{Satiated}(\mathit{Cat}), \mathit{Tail}(\mathit{Tail}(P_1)))
\end{gathered} \tag{11.59}
$$

From (11.57), (11.58), and (11.55), we have

$$
\begin{gathered}
\neg \exists a, g, p\, (e = \mathit{AddPlan}(a, g, p) \land \\
\neg \mathit{HoldsAt}(\mathit{Believe}(a, g), 2) \land \\
\mathit{HoldsAt}(\mathit{Goal}(a, g), 2) \land \\
\mathit{SelectedPlan}(a, g, p, 2) \land \\
\neg \exists p_1\, \mathit{HoldsAt}(\mathit{Plan}(a, g, p_1), 2))
\end{gathered} \tag{11.60}
$$

From (11.54) and (11.58), we have

$$
\begin{gathered}
\neg \exists e, a, g\, (e = \mathit{DropGoal}(a, g) \land \\
\mathit{HoldsAt}(\mathit{Goal}(a, g), 2) \land \\
\mathit{HoldsAt}(\mathit{Believe}(a, g), 2))
\end{gathered} \tag{11.61}
$$

From (11.52), we have

$$
\begin{gathered}
\neg \exists e, a, f, s\, (e = \mathit{Eat}(a, f) \land \\
\neg \mathit{HoldsAt}(\mathit{Satiated}(a), 2) \land \\
\mathit{HoldsAt}(\mathit{On}(a, s), 2) \land \\
\mathit{HoldsAt}(\mathit{On}(f, s), 2))
\end{gathered} \tag{11.62}
$$

From (11.53), (11.15), L5, and L6, we have

$$SoundPlan(Cat, Satiated(Cat), Tail(P_1), 2)$$

From this, (11.60), (11.61), (11.62), (11.56), (11.55), (11.54), (11.29), L1, L2, L3, and L4, we have (11.59).

Timepoint 3: Fluents Using similar arguments we can show

$$\begin{gathered} HoldsAt(On(o,s),3) \Leftrightarrow \\ (o = Cat \wedge s = Chair) \vee \\ (o = Food1 \wedge s = Table) \vee \\ (o = Food2 \wedge s = Shelf\,) \end{gathered} \tag{11.63}$$

$$\begin{gathered} HoldsAt(Believe(a, CanJump(s_1, s_2)), 3) \Leftrightarrow \\ (a = Cat \wedge s_1 = Floor \wedge s_2 = Chair) \vee \\ (a = Cat \wedge s_1 = Chair \wedge s_2 = Shelf\,) \vee \\ (a = Cat \wedge s_1 = Shelf \wedge s_2 = Table) \end{gathered} \tag{11.64}$$

$$\neg HoldsAt(Believe(Cat, Satiated(Cat)), 3) \tag{11.65}$$

$$HoldsAt(Plan(Cat, Satiated(Cat), Tail(Tail(P_1))), 3) \tag{11.66}$$

$$\begin{gathered} a \neq Cat \vee g \neq Satiated(Cat) \vee p \neq Tail(Tail(P_1)) \Rightarrow \\ \neg HoldsAt(Plan(a, g, p), 3) \end{gathered} \tag{11.67}$$

$$HoldsAt(Goal(Cat, Satiated(Cat)), 3) \tag{11.68}$$

$$a \neq Cat \vee g \neq Satiated(Cat) \Rightarrow \neg HoldsAt(Goal(a, g), 3) \tag{11.69}$$

Timepoint 3: Events At timepoint 3, the cat no longer believes that the plan is sound. So the cat does not act on the plan, and the plan is dropped. We proceed to show

$$\begin{gathered} Happens(e, 3) \Leftrightarrow \\ e = DropPlan(Cat, Satiated(Cat), Tail(Tail(P_1))) \end{gathered} \tag{11.70}$$

From (11.68), (11.69), and (11.66), we have

$$\begin{gathered} \neg\exists a, g, p\ (e = AddPlan(a, g, p) \wedge \\ \neg HoldsAt(Believe(a, g), 3) \wedge \\ HoldsAt(Goal(a, g), 3) \wedge \\ SelectedPlan(a, g, p, 3) \wedge \\ \neg\exists p_1\ HoldsAt(Plan(a, g, p_1), 3)) \end{gathered} \tag{11.71}$$

From (11.64), (11.15), L5, and L6, we have

$$\neg SoundPlan(Cat, Satiated(Cat), Tail(Tail(P_1)), 3)$$

From this, (11.66), and (11.67), we have

$$\begin{aligned}\neg\exists e, a, g, p\,(e = Head(p)\ \wedge\\ HoldsAt(Plan(a, g, p), 3)\ \wedge\\ \neg HoldsAt(Believe(a, g), 3)\ \wedge\\ SoundPlan(a, g, p, 3))\end{aligned} \tag{11.72}$$

$$\begin{aligned}\neg\exists e, a, g, p\,(e = AddPlan(a, g, Tail(p))\ \wedge\\ HoldsAt(Plan(a, g, p), 3)\ \wedge\\ Length(p) > 1\ \wedge\\ \neg HoldsAt(Believe(a, g), 3)\ \wedge\\ SoundPlan(a, g, p, 3))\end{aligned} \tag{11.73}$$

From (11.65) and (11.69), we have

$$\begin{aligned}\neg\exists e, a, g\,(e = DropGoal(a, g)\ \wedge\\ HoldsAt(Goal(a, g), 3)\ \wedge\\ HoldsAt(Believe(a, g), 3))\end{aligned} \tag{11.74}$$

From (11.63), we have

$$\begin{aligned}\neg\exists e, a, f, s\,(e = Eat(a, f)\ \wedge\\ \neg HoldsAt(Satiated(a), 3)\ \wedge\\ HoldsAt(On(a, s), 3)\ \wedge\\ HoldsAt(On(f, s), 3))\end{aligned} \tag{11.75}$$

From (11.71), (11.72), (11.73), (11.74), (11.75), (11.67), (11.66), and (11.29), we have (11.70).

Timepoint 4: Fluents Using similar arguments we can show

$$\begin{aligned}HoldsAt(On(o, s), 4) \Leftrightarrow\\ (o = Cat \wedge s = Chair) \vee\\ (o = Food1 \wedge s = Table) \vee\\ (o = Food2 \wedge s = Shelf\,)\end{aligned} \tag{11.76}$$

$$\begin{aligned}HoldsAt(Believe(a, On(o, s)), 4) \Leftrightarrow\\ (a = Cat \wedge o = Cat \wedge s = Chair) \vee\\ (a = Cat \wedge o = Food1 \wedge s = Table)\end{aligned} \tag{11.77}$$

$$\begin{aligned}HoldsAt(Believe(a, CanJump(s_1, s_2)), 4) \Leftrightarrow\\ (a = Cat \wedge s_1 = Floor \wedge s_2 = Chair) \vee\\ (a = Cat \wedge s_1 = Chair \wedge s_2 = Shelf\,) \vee\\ (a = Cat \wedge s_1 = Shelf \wedge s_2 = Table)\end{aligned} \tag{11.78}$$

$$\neg HoldsAt(Believe(Cat, Satiated(Cat)), 4) \tag{11.79}$$

$$\neg HoldsAt(Plan(a,g,p),4) \tag{11.80}$$

$$HoldsAt(Goal(Cat,Satiated(Cat)),4) \tag{11.81}$$

$$a \neq Cat \vee g \neq Satiated(Cat) \Rightarrow \neg HoldsAt(Goal(a,g),4) \tag{11.82}$$

Timepoint 4: Events Because no plans are active at timepoint 4, a new plan is activated. Let $P_2 = [Jump(Cat, Chair, Shelf), Wait(Cat), Jump(Cat, Shelf, Table), Eat(Cat, Food1)]$. We proceed to show

$$\begin{aligned} &Happens(e,4) \Leftrightarrow \\ &e = AddPlan(Cat, Satiated(Cat), P_2) \end{aligned} \tag{11.83}$$

From (11.80), we have

$$\begin{aligned} &\neg\exists e,a,g,p\,(e = Head(p) \wedge \\ &HoldsAt(Plan(a,g,p),4) \wedge \\ &\neg HoldsAt(Believe(a,g),4) \wedge \\ &SoundPlan(a,g,p,4)) \end{aligned} \tag{11.84}$$

$$\neg\exists e,a,g,p\,(e = DropPlan(a,g,p) \wedge HoldsAt(Plan(a,g,p),4)) \tag{11.85}$$

$$\begin{aligned} &\neg\exists e,a,g,p\,(e = AddPlan(a,g,Tail(p)) \wedge \\ &HoldsAt(Plan(a,g,p),4) \wedge \\ &Length(p) > 1 \wedge \\ &\neg HoldsAt(Believe(a,g),4) \wedge \\ &SoundPlan(a,g,p,4)) \end{aligned} \tag{11.86}$$

From (11.81), (11.82), and (11.79), we have

$$\begin{aligned} &\neg\exists e,a,g\,(e = DropGoal(a,g) \wedge \\ &HoldsAt(Goal(a,g),4) \wedge \\ &HoldsAt(Believe(a,g),4)) \end{aligned} \tag{11.87}$$

From (11.76), we have

$$\begin{aligned} &\neg\exists e,a,f,s\,(e = Eat(a,f) \wedge \\ &\neg HoldsAt(Satiated(a),4) \wedge \\ &HoldsAt(On(a,s),4) \wedge \\ &HoldsAt(On(f,s),4)) \end{aligned} \tag{11.88}$$

From (11.77), (11.78), and (11.14), we have

$$SelectedPlan(a,g,p,4) \Leftrightarrow a = Cat \wedge g = Satiated(Cat) \wedge p = P_2$$

From this, (11.81), (11.82), (11.80), (11.79), (11.84), (11.85), (11.86), (11.87), (11.88), and (11.29), we have (11.83).

Timepoint 5: Fluents Using similar arguments we can show

$$\begin{aligned} & HoldsAt(On(o,s),5) \Leftrightarrow \\ & (o = Cat \land s = Chair) \lor \\ & (o = Food1 \land s = Table) \lor \\ & (o = Food2 \land s = Shelf\,) \end{aligned} \tag{11.89}$$

$$\begin{aligned} & HoldsAt(Believe(a, CanJump(s_1,s_2)),5) \Leftrightarrow \\ & (a = Cat \land s_1 = Floor \land s_2 = Chair) \lor \\ & (a = Cat \land s_1 = Chair \land s_2 = Shelf\,) \lor \\ & (a = Cat \land s_1 = Shelf \land s_2 = Table) \end{aligned} \tag{11.90}$$

$$\neg HoldsAt(Believe(Cat, Satiated(Cat)),5) \tag{11.91}$$

$$HoldsAt(Plan(Cat, Satiated(Cat), P_2),5) \tag{11.92}$$

$$\begin{aligned} & a \neq Cat \lor g \neq Satiated(Cat) \lor p \neq P_2 \Rightarrow \\ & \neg HoldsAt(Plan(a,g,p),5) \end{aligned} \tag{11.93}$$

$$HoldsAt(Goal(Cat, Satiated(Cat)),5) \tag{11.94}$$

$$a \neq Cat \lor g \neq Satiated(Cat) \Rightarrow \neg HoldsAt(Goal(a,g),4) \tag{11.95}$$

Timepoint 5: Events The new plan is acted on. We proceed to show

$$\begin{aligned} & Happens(e,5) \Leftrightarrow \\ & e = Jump(Cat, Chair, Shelf\,) \lor \\ & e = DropPlan(Cat, Satiated(Cat), P_2) \lor \\ & e = AddPlan(Cat, Satiated(Cat), Tail(P_2)) \end{aligned} \tag{11.96}$$

From (11.94), (11.95), and (11.92), we have

$$\begin{aligned} & \neg\exists a,g,p\,(e = AddPlan(a,g,p) \land \\ & \neg HoldsAt(Believe(a,g),5) \land \\ & HoldsAt(Goal(a,g),5) \land \\ & SelectedPlan(a,g,p,5) \land \\ & \neg\exists p_1\; HoldsAt(Plan(a,g,p_1),5)) \end{aligned} \tag{11.97}$$

From (11.91) and (11.95), we have

$$\begin{aligned} & \neg\exists e,a,g\,(e = DropGoal(a,g) \land \\ & HoldsAt(Goal(a,g),5) \land \\ & HoldsAt(Believe(a,g),5)) \end{aligned} \tag{11.98}$$

From (11.89), we have

$$\begin{aligned} & \neg\exists e,a,f,s\,(e = Eat(a,f) \land \\ & \neg HoldsAt(Satiated(a),5) \land \\ & HoldsAt(On(a,s),5) \land \\ & HoldsAt(On(f,s),5)) \end{aligned} \tag{11.99}$$

From (11.90), (11.15), L5, and L6, we have

$$SoundPlan(Cat, Satiated(Cat), P_2, 5)$$

From this, (11.97), (11.98), (11.99), (11.93), (11.92), (11.91), (11.29), L1, L3, and L4, we have (11.96).

Timepoint 6: Fluents Using similar arguments we can show

$$\begin{aligned} & HoldsAt(On(o, s), 6) \Leftrightarrow \\ & (o = Cat \wedge s = Shelf\,) \vee \\ & (o = Food1 \wedge s = Table) \vee \\ & (o = Food2 \wedge s = Shelf\,) \end{aligned} \tag{11.100}$$

$$\neg HoldsAt(Believe(Cat, Satiated(Cat)), 6) \tag{11.101}$$

Timepoint 6: Events From (11.101), (11.100), (11.16) (which follows from (11.29)), we have $Happens(Eat(Cat, Food2), 6)$ as required. (At timepoint 7, the cat is satiated, the plan is not further acted upon, and the goal to be satiated is dropped.) ■

11.2 EMOTIONS

We use the theory of human emotions and emotional reactions to events proposed by Andrew Ortony, Gerald L. Clore, and Allan Collins.

11.2.1 EMOTION THEORY

The emotion theory specifies several emotion classes, each consisting of several emotions or emotional reactions. For each emotion, *eliciting conditions* determine under what circumstances the emotion is elicited.

> *Well-being emotions* are reactions to desirable or undesirable events. For example, if a desirable event occurs, then a *joy* emotion is elicited, whereas if an undesirable event occurs, then a *distress* emotion is elicited.
> *Fortunes-of-others emotions* are reactions to events involving others. They vary according to whether they are desirable or undesirable for the self and whether they are believed to be desirable or undesirable for the other person. For example, if an event occurs that is undesirable for the self and desirable for the other person, then a *resentment* emotion is elicited.
> *Prospect-based emotions* are reactions to events that are likely to occur but have not yet occurred and reactions to anticipated events that occur or fail to occur. They vary according to whether they are desirable or undesirable and whether they have not yet occurred, they have occurred, or they have failed to occur. For example, if a desirable event is likely to occur, then a *hope* emotion is elicited; if an undesirable event has failed to occur, then the *relief* emotion is elicited.
> *Attribution emotions* are reactions to actions performed by the self or others of which the self approves or disapproves. For example, if the self performs an

action approved by the self, then the *pride* emotion is elicited; if another person performs an action approved by the self, then the *appreciation* emotion is elicited.
Compound emotions are combinations of well-being and attribution emotions. For example, *gratitude* is a combination of joy and appreciation.
Attraction emotions are positive and negative reactions to objects. For example, a person *likes* or *dislikes* various objects.

11.2.2 FORMALIZATION

We now formalize in the event calculus (parts of) the theory of Ortony, Clore, and Collins. We have the following sorts:

- a belief sort, with variables $b, b_1, b_2, \ldots$
- an event sort, which is a subsort of the belief sort
- a negated event sort, which is also a subsort of the belief sort
- a fluent sort, with variables $f, f_1, f_2, \ldots$
- an object sort, with variables $o, o_1, o_2, \ldots$
- an agent sort, which is a subsort of the object sort, with variables $a, a_1, a_2, \ldots$
- a real number sort, with variables $x, x_1, x_2, \ldots$
- an event sort, with variables $e, e_1, e_2, \ldots$

We also introduce a function symbol *Not*, which denotes a function that maps events to their negations.

We start with the following fluents that represent factors used to specify the eliciting conditions of emotions.

$Believe(a, b)$: If b is an event term, then agent a believes that b has occurred. If b is a negated event term $Not(e)$, then agent a believes that e has not occurred.
$Desirability(a_1, a_2, e, x)$: Agent a_1 believes that the desirability of event e to agent a_2 is x, where $-1 \leq x \leq 1$. Agents a_1 and a_2 may be the same.
$Praiseworthiness(a_1, a_2, e, x)$: Agent a_1 believes that the praiseworthiness of an action e performed by agent a_2 is x, where $-1 \leq x \leq 1$. Agents a_1 and a_2 may be the same.
$Anticipate(a, e, x)$: Agent a anticipates that event e will occur with likelihood x, where $0 \leq x \leq 1$.

We have the following fluents for well-being emotions:

$Joy(a, e)$: Agent a is joyful about event e.
$Distress(a, e)$: Agent a is distressed about event e.

We have the following fluents for fortunes-of-others emotions:

$HappyFor(a_1, a_2, e)$: Agent a_1 is happy for agent a_2 regarding event e.
$SorryFor(a_1, a_2, e)$: Agent a_1 is sorry for agent a_2 regarding event e.

Resentment(a_1, a_2, e): Agent a_1 is resentful of agent a_2 regarding event e.
Gloating(a_1, a_2, e): Agent a_1 gloats toward agent a_2 regarding event e.

We have the following fluents for prospect-based emotions:

Hope(a, e): Agent a is hopeful about an anticipated event e.
Fear(a, e): Agent a is fearful about an anticipated event e.
Satisfaction(a, e): Agent a is satisfied that event e occurred.
Disappointment(a, e): Agent a is disappointed that event e did not occur.
Relief(a, e): Agent a is relieved that event e did not occur.
FearsConfirmed(a, e): The fears of a are confirmed by the occurrence of event e.

We have the following fluents for attribution emotions:

Pride(a, e): Agent a is proud of performing action e.
SelfReproach(a, e): Agent a feels self-reproach for performing action e.
Appreciation(a_1, a_2, e): Agent a_1 is appreciative of agent a_2 for performing action e.
Reproach(a_1, a_2, e): Agent a_1 is reproachful of agent a_2 for performing action e.

We have the following fluents for compound emotions:

Gratitude(a_1, a_2, e): Agent a_1 is grateful toward agent a_2 regarding event e.
Anger(a_1, a_2, e): Agent a_1 is angry at agent a_2 regarding event e.
Gratification(a, e): Agent a is gratified about event e.
Remorse(a, e): Agent a is remorseful about event e.

We have the following fluents for attraction emotions:

Like(a, o): Agent a likes object o.
Dislike(a, o): Agent a dislikes object o.

We now present a number of axioms, starting with those for adding and dropping these fluents.

Axiom E1.

$$Initiates(AddBelief(a, b), Believe(a, b), t)$$

Axiom E2.

$$Terminates(DropBelief(a, b), Believe(a, b), t)$$

Axiom E3.

$$Initiates(AddDesirability(a_1, a_2, e, x), Desirability(a_1, a_2, e, x), t)$$

Axiom E4.

$$Terminates(DropDesirability(a_1, a_2, e, x), Desirability(a_1, a_2, e, x), t)$$

Axiom E5.

$$Initiates(AddPraiseworthiness(a_1, a_2, e, x), Praiseworthiness(a_1, a_2, e, x), t)$$

Axiom E6.

$$Terminates(DropPraiseworthiness(a_1, a_2, e, x), Praiseworthiness(a_1, a_2, e, x), t)$$

Axiom E7.

$$Initiates(AddAnticipate(a, e, x), Anticipate(a, e, x), t)$$

Axiom E8.

$$Terminates(DropAnticipate(a, e, x), Anticipate(a, e, x), t)$$

Axiom E9.

$$Initiates(AddJoy(a, e), Joy(a, e), t)$$

Axiom E10.

$$Terminates(DropJoy(a, e), Joy(a, e), t)$$

Axiom E11.

$$Initiates(AddDistress(a, e), Distress(a, e), t)$$

Axiom E12.

$$Terminates(DropDistress(a, e), Distress(a, e), t)$$

Axiom E13.

$$Initiates(AddHappyFor(a_1, a_2, e), HappyFor(a_1, a_2, e), t)$$

Axiom E14.

$$Terminates(DropHappyFor(a_1, a_2, e), HappyFor(a_1, a_2, e), t)$$

Axiom E15.

$$Initiates(AddSorryFor(a_1, a_2, e), SorryFor(a_1, a_2, e), t)$$

Axiom E16.

$$Terminates(DropSorryFor(a_1, a_2, e), SorryFor(a_1, a_2, e), t)$$

Axiom E17.

$$Initiates(AddResentment(a_1, a_2, e), Resentment(a_1, a_2, e), t)$$

Axiom E18.

$$Terminates(DropResentment(a_1, a_2, e), Resentment(a_1, a_2, e), t)$$

Axiom E19.

$$Initiates(AddGloating(a_1, a_2, e), Gloating(a_1, a_2, e), t)$$

Axiom E20.

$$Terminates(DropGloating(a_1, a_2, e), Gloating(a_1, a_2, e), t)$$

Axiom E21.

$$Initiates(AddHope(a, e), Hope(a, e), t)$$

Axiom E22.

$$Terminates(DropHope(a, e), Hope(a, e), t)$$

Axiom E23.

$$Initiates(AddFear(a, e), Fear(a, e), t)$$

Axiom E24.

$$Terminates(DropFear(a, e), Fear(a, e), t)$$

Axiom E25.

$$Initiates(AddSatisfaction(a, e), Satisfaction(a, e), t)$$

Axiom E26.

$$Terminates(DropSatisfaction(a, e), Satisfaction(a, e), t)$$

Axiom E27.

$$Initiates(AddDisappointment(a, e), Disappointment(a, e), t)$$

Axiom E28.

$$Terminates(DropDisappointment(a, e), Disappointment(a, e), t)$$

Axiom E29.

$$Initiates(AddRelief(a, e), Relief(a, e), t)$$

Axiom E30.

$$Terminates(DropRelief(a, e), Relief(a, e), t)$$

Axiom E31.

$$Initiates(AddFearsConfirmed(a, e), FearsConfirmed(a, e), t)$$

Axiom E32.

$$Terminates(DropFearsConfirmed(a, e), FearsConfirmed(a, e), t)$$

Axiom E33.

$$Initiates(AddPride(a, e), Pride(a, e), t)$$

Axiom E34.

$$Terminates(DropPride(a, e), Pride(a, e), t)$$

Axiom E35.

$$Initiates(AddSelfReproach(a, e), SelfReproach(a, e), t)$$

Axiom E36.

$$Terminates(DropSelfReproach(a, e), SelfReproach(a, e), t)$$

Axiom E37.

$$Initiates(AddAppreciation(a_1, a_2, e), Appreciation(a_1, a_2, e), t)$$

Axiom E38.

$$Terminates(DropAppreciation(a_1, a_2, e), Appreciation(a_1, a_2, e), t)$$

Axiom E39.

$$Initiates(AddReproach(a_1, a_2, e), Reproach(a_1, a_2, e), t)$$

Axiom E40.

$$Terminates(DropReproach(a_1, a_2, e), Reproach(a_1, a_2, e), t)$$

Axiom E41.

$$Initiates(AddGratitude(a_1, a_2, e), Gratitude(a_1, a_2, e), t)$$

Axiom E42.

$$Terminates(DropGratitude(a_1, a_2, e), Gratitude(a_1, a_2, e), t)$$

Axiom E43.

$$Initiates(AddAnger(a_1, a_2, e), Anger(a_1, a_2, e), t)$$

Axiom E44.

$$Terminates(DropAnger(a_1, a_2, e), Anger(a_1, a_2, e), t)$$

Axiom E45.

$$Initiates(AddGratification(a, e), Gratification(a, e), t)$$

Axiom E46.

$$Terminates(DropGratification(a, e), Gratification(a, e), t)$$

Axiom E47.

$$Initiates(AddRemorse(a, e), Remorse(a, e), t)$$

Axiom E48.

$$Terminates(DropRemorse(a, e), Remorse(a, e), t)$$

Axiom E49.

$$Initiates(AddLike(a, o), Like(a, o), t)$$

Axiom E50.

$$Terminates(DropLike(a, o), Like(a, o), t)$$

Axiom E51.

$$Initiates(AddDislike(a, o), Dislike(a, o), t)$$

Axiom E52.

$$Terminates(DropDislike(a, o), Dislike(a, o), t)$$

We continue with state constraints on desirability, praiseworthiness, and anticipation.

Axiom E53.

$$\begin{gathered} HoldsAt(Desirability(a_1, a_2, e, x_1), t) \wedge \\ HoldsAt(Desirability(a_1, a_2, e, x_2), t) \Rightarrow \\ x_1 = x_2 \end{gathered}$$

Axiom E54.

$$HoldsAt(Desirability(a_1, a_2, e, x), t) \Rightarrow -1 \leq x \leq 1$$

Axiom E55.

$$HoldsAt(Praiseworthiness(a_1, a_2, e, x_1), t) \wedge$$
$$HoldsAt(Praiseworthiness(a_1, a_2, e, x_2), t) \Rightarrow$$
$$x_1 = x_2$$

Axiom E56.

$$HoldsAt(Praiseworthiness(a_1, a_2, e, x), t) \Rightarrow -1 \leq x \leq 1$$

Axiom E57.

$$HoldsAt(Anticipate(a, e, x_1), t) \wedge$$
$$HoldsAt(Anticipate(a, e, x_2), t) \Rightarrow$$
$$x_1 = x_2$$

Axiom E58.

$$HoldsAt(Anticipate(a, e, x), t) \Rightarrow 0 \leq x \leq 1$$

We now present axioms for eliciting emotions.

Axiom E59.

$$\neg HoldsAt(Joy(a, e), t) \wedge$$
$$HoldsAt(Desirability(a, a, e, x), t) \wedge x > 0 \wedge$$
$$HoldsAt(Believe(a, e), t) \Rightarrow$$
$$Happens(AddJoy(a, e), t)$$

Axiom E60.

$$\neg HoldsAt(Distress(a, e), t) \wedge$$
$$HoldsAt(Desirability(a, a, e, x), t) \wedge x < 0 \wedge$$
$$HoldsAt(Believe(a, e), t) \Rightarrow$$
$$Happens(AddDistress(a, e), t)$$

Axiom E61.

$$\neg HoldsAt(HappyFor(a_1, a_2, e), t) \wedge$$
$$HoldsAt(Desirability(a_1, a_2, e, x_1), t) \wedge x_1 > 0 \wedge$$
$$HoldsAt(Desirability(a_1, a_1, e, x_2), t) \wedge x_2 > 0 \wedge$$
$$HoldsAt(Like(a_1, a_2), t) \wedge$$
$$HoldsAt(Believe(a_1, e), t) \wedge$$
$$a_1 \neq a_2 \Rightarrow$$
$$Happens(AddHappyFor(a_1, a_2, e), t)$$

Axiom E62.

$$\neg HoldsAt(SorryFor(a_1, a_2, e), t) \wedge$$
$$HoldsAt(Desirability(a_1, a_2, e, x_1), t) \wedge x_1 < 0 \wedge$$
$$HoldsAt(Desirability(a_1, a_1, e, x_2), t) \wedge x_2 < 0 \wedge$$
$$HoldsAt(Like(a_1, a_2), t) \wedge$$
$$HoldsAt(Believe(a_1, e), t) \wedge$$

$$a_1 \neq a_2 \Rightarrow$$
$$Happens(AddSorryFor(a_1, a_2, e), t)$$

Axiom E63.

$$\neg HoldsAt(Resentment(a_1, a_2, e), t) \wedge$$
$$HoldsAt(Desirability(a_1, a_2, e, x_1), t) \wedge x_1 > 0 \wedge$$
$$HoldsAt(Desirability(a_1, a_1, e, x_2), t) \wedge x_2 < 0 \wedge$$
$$HoldsAt(Dislike(a_1, a_2), t) \wedge$$
$$HoldsAt(Believe(a_1, e), t) \wedge$$
$$a_1 \neq a_2 \Rightarrow$$
$$Happens(AddResentment(a_1, a_2, e), t)$$

Axiom E64.

$$\neg HoldsAt(Gloating(a_1, a_2, e), t) \wedge$$
$$HoldsAt(Desirability(a_1, a_2, e, x_1), t) \wedge x_1 < 0 \wedge$$
$$HoldsAt(Desirability(a_1, a_1, e, x_2), t) \wedge x_2 > 0 \wedge$$
$$HoldsAt(Dislike(a_1, a_2), t) \wedge$$
$$HoldsAt(Believe(a_1, e), t) \wedge$$
$$a_1 \neq a_2 \Rightarrow$$
$$Happens(AddGloating(a_1, a_2, e), t)$$

Axiom E65.

$$\neg HoldsAt(Hope(a, e), t) \wedge$$
$$HoldsAt(Desirability(a, a, e, x_1), t) \wedge x_1 > 0 \wedge$$
$$HoldsAt(Anticipate(a, e, x_2), t) \wedge x_2 > 0 \wedge$$
$$a_1 \neq a_2 \Rightarrow$$
$$Happens(AddHope(a, e), t)$$

Axiom E66.

$$\neg HoldsAt(Fear(a, e), t) \wedge$$
$$HoldsAt(Desirability(a, a, e, x_1), t) \wedge x_1 < 0 \wedge$$
$$HoldsAt(Anticipate(a, e, x_2), t) \wedge x_2 > 0 \Rightarrow$$
$$Happens(AddFear(a, e), t)$$

Axiom E67.

$$\neg HoldsAt(Satisfaction(a, e), t) \wedge$$
$$HoldsAt(Desirability(a, a, e, x_1), t) \wedge x_1 > 0 \wedge$$
$$HoldsAt(Anticipate(a, e, x_2), t) \wedge x_2 > 0 \wedge$$
$$HoldsAt(Believe(a, e), t) \Rightarrow$$
$$Happens(AddSatisfaction(a, e), t)$$

Axiom E68.

$$\neg HoldsAt(Disappointment(a, e), t) \wedge$$
$$HoldsAt(Desirability(a, a, e, x_1), t) \wedge x_1 > 0 \wedge$$

$$\begin{aligned}&HoldsAt(Anticipate(a,e,x_2),t) \wedge x_2 > 0 \wedge\\&HoldsAt(Believe(a,\neg e),t) \Rightarrow\\&Happens(AddDisappointment(a,e),t)\end{aligned}$$

Axiom E69.

$$\begin{aligned}&\neg HoldsAt(Relief(a,e),t) \wedge\\&HoldsAt(Desirability(a,a,e,x_1),t) \wedge x_1 < 0 \wedge\\&HoldsAt(Anticipate(a,e,x_2),t) \wedge x_2 > 0 \wedge\\&HoldsAt(Believe(a,\neg e),t) \Rightarrow\\&Happens(AddRelief(a,e),t)\end{aligned}$$

Axiom E70.

$$\begin{aligned}&\neg HoldsAt(FearsConfirmed(a,e),t) \wedge\\&HoldsAt(Desirability(a,a,e,x_1),t) \wedge x_1 < 0 \wedge\\&HoldsAt(Anticipate(a,e,x_2),t) \wedge x_2 > 0 \wedge\\&HoldsAt(Believe(a,e),t) \Rightarrow\\&Happens(AddFearsConfirmed(a,e),t)\end{aligned}$$

Axiom E71.

$$\begin{aligned}&\neg HoldsAt(Pride(a,e),t) \wedge\\&HoldsAt(Praiseworthiness(a,a,e,x),t) \wedge x > 0 \wedge\\&HoldsAt(Believe(a,e),t) \Rightarrow\\&Happens(AddPride(a,e),t)\end{aligned}$$

Axiom E72.

$$\begin{aligned}&\neg HoldsAt(SelfReproach(a,e),t) \wedge\\&HoldsAt(Praiseworthiness(a,a,e,x),t) \wedge x < 0 \wedge\\&HoldsAt(Believe(a,e),t) \Rightarrow\\&Happens(AddSelfReproach(a,e),t)\end{aligned}$$

Axiom E73.

$$\begin{aligned}&\neg HoldsAt(Appreciation(a_1,a_2,e),t) \wedge\\&HoldsAt(Praiseworthiness(a_1,a_2,e,x),t) \wedge x > 0 \wedge\\&HoldsAt(Believe(a_1,e),t) \wedge a_1 \neq a_2 \Rightarrow\\&Happens(AddAppreciation(a_1,a_2,e),t)\end{aligned}$$

Axiom E74.

$$\begin{aligned}&\neg HoldsAt(Reproach(a_1,a_2,e),t) \wedge\\&HoldsAt(Praiseworthiness(a_1,a_2,e,x),t) \wedge x < 0 \wedge\\&HoldsAt(Believe(a_1,e),t) \wedge a_1 \neq a_2 \Rightarrow\\&Happens(AddReproach(a_1,a_2,e),t)\end{aligned}$$

We specify compound emotions using state constraints.

Axiom E75.

$$\begin{gathered}HoldsAt(Gratitude(a_1, a_2, e), t) \Leftrightarrow \\ HoldsAt(Joy(a_1, e), t) \wedge \\ HoldsAt(Appreciation(a_1, a_2, e), t)\end{gathered}$$

Axiom E76.

$$\begin{gathered}HoldsAt(Anger(a_1, a_2, e), t) \Leftrightarrow \\ HoldsAt(Distress(a_1, e), t) \wedge \\ HoldsAt(Reproach(a_1, a_2, e), t)\end{gathered}$$

Axiom E77.

$$\begin{gathered}HoldsAt(Gratification(a, e), t) \Leftrightarrow \\ HoldsAt(Joy(a, e), t) \wedge \\ HoldsAt(Pride(a, e), t)\end{gathered}$$

Axiom E78.

$$\begin{gathered}HoldsAt(Remorse(a, e), t) \Leftrightarrow \\ HoldsAt(Distress(a, e), t) \wedge \\ HoldsAt(SelfReproach(a, e), t)\end{gathered}$$

11.2.3 EXAMPLE: LOTTERY

If we are told that Lisa likes Kate, who just won the lottery, we should infer that Kate will be happy and Lisa will be happy for Kate. We assume that when any agent wins the lottery, all agents are aware of it:

$$Initiates(WinLottery(a_1), Believe(a_2, WinLottery(a_1)), t) \tag{11.102}$$

We have the following observations and narrative. Initially, Kate is not joyful about any events, and Lisa is not happy for Kate regarding any events:

$$\neg ReleasedAt(f, t) \tag{11.103}$$

$$\neg HoldsAt(Joy(Kate, e), 0) \tag{11.104}$$

$$\neg HoldsAt(HappyFor(Lisa, Kate, e), 0) \tag{11.105}$$

Kate and Lisa do not believe any events have occurred:

$$\neg HoldsAt(Believe(Kate, e), 0) \tag{11.106}$$

$$\neg HoldsAt(Believe(Lisa, e), 0) \tag{11.107}$$

Lisa likes Kate, Kate believes that her winning the lottery is highly desirable to herself, Lisa believes that Kate winning the lottery is highly desirable to Kate, and Lisa believes that Kate winning the lottery is highly desirable to Lisa:

$$HoldsAt(Like(Lisa, Kate), t) \tag{11.108}$$

$$HoldsAt(Desirability(Kate, Kate, WinLottery(Kate), 1), t) \tag{11.109}$$

$$HoldsAt(Desirability(Lisa, Kate, WinLottery(Kate), 1), t) \tag{11.110}$$

$$HoldsAt(Desirability(Lisa, Lisa, WinLottery(Kate), 1), t) \tag{11.111}$$

Kate wins the lottery:

$$Happens(WinLottery(Kate), 0) \tag{11.112}$$

We can show that Kate will be happy, and Lisa will be happy for Kate.

Proposition 11.2. *Let Σ be the conjunction of E9, E13, and (11.102). Let Δ be the conjunction of E59, E61, and (11.112). Let Ω be $U[AddJoy, AddHappyFor, WinLottery] \wedge U[Joy, HappyFor, Believe, Desirability, Like]$. Let Ψ be the conjunction of E53 and E54. Let Γ be the conjunction of (11.103), (11.104), (11.105), (11.106), (11.107), (11.108), (11.109), (11.110), and (11.111). Then we have*

$$\begin{aligned}
&CIRC[\Sigma; Initiates, Terminates, Releases] \wedge CIRC[\Delta; Happens] \wedge \\
&\Omega \wedge \Psi \wedge \Gamma \wedge \mathrm{DEC} \vdash HoldsAt(Joy(Kate, WinLottery(Kate)), 2) \wedge \\
&HoldsAt(HappyFor(Lisa, Kate, WinLottery(Kate)), 2)
\end{aligned}$$

Proof. From $CIRC[\Sigma; Initiates, Terminates, Releases]$ and Theorems 2.2 and 2.1, we have

$$\begin{aligned}
&Initiates(e, f, t) \Leftrightarrow \\
&\exists a, e_1\ (e = AddJoy(a, e_1) \wedge f = Joy(a, e_1)) \vee \\
&\exists a_1, a_2, e_1\ (e = AddHappyFor(a_1, a_2, e_1) \wedge f = HappyFor(a_1, a_2, e_1)) \vee \\
&\exists a_1, a_2\ (e = WinLottery(a_1) \wedge f = Believe(a_2, WinLottery(a_1)))
\end{aligned} \tag{11.113}$$

$$\neg Terminates(e, f, t) \tag{11.114}$$

$$\neg Releases(e, f, t) \tag{11.115}$$

From $CIRC[\Delta; Happens]$ and Theorem 2.1, we have

$$\begin{aligned}
&Happens(e, t) \Leftrightarrow \\
&\exists a, e_1, x\ (e = AddJoy(a, e_1) \wedge \\
&\neg HoldsAt(Joy(a, e_1), t) \wedge \\
&HoldsAt(Desirability(a, a, e_1, x), t) \wedge \\
&x > 0 \wedge \\
&HoldsAt(Believe(a, e_1), t)) \vee \\
&\exists a_1, a_2, e_1, x_1, x_2\ (e = AddHappyFor(a_1, a_2, e_1) \wedge \\
&\neg HoldsAt(HappyFor(a_1, a_2, e_1), t) \wedge \\
&HoldsAt(Desirability(a_1, a_2, e_1, x_1), t) \wedge \\
&x_1 > 0 \wedge \\
&HoldsAt(Desirability(a_1, a_1, e_1, x_2), t) \wedge \\
&x_2 > 0 \wedge \\
&HoldsAt(Like(a_1, a_2), t) \wedge \\
&HoldsAt(Believe(a_1, e_1), t) \wedge \\
&a_1 \neq a_2) \vee \\
&(e = WinLottery(Kate) \wedge t = 0)
\end{aligned} \tag{11.116}$$

From (11.112) (which follows from (11.116)), (11.102) (which follows from (11.113)), and DEC9, we have the following:

$$HoldsAt(Believe(Kate, WinLottery(Kate)), 1) \quad (11.117)$$

$$HoldsAt(Believe(Lisa, WinLottery(Kate)), 1) \quad (11.118)$$

From (11.113), (11.116), and (11.106), we have $\neg\exists e\,(Happens(e, 0) \wedge Initiates(e, Joy(Kate, e_1), 0))$. From this, (11.104), (11.103), and DEC6, we have $\neg HoldsAt(Joy(Kate, e), 1)$. From this, (11.109), $1 > 0$, (11.117), and E59 (which follows from (11.116)), we have

$$Happens(AddJoy(Kate, WinLottery(Kate)), 1) \quad (11.119)$$

From (11.113), (11.116), and (11.107), we have $\neg\exists e\,(Happens(e, 0) \wedge Initiates(e, HappyFor(Lisa, Kate, e_1), 0))$. From this, (11.105), (11.103), and DEC6, we have $\neg HoldsAt(HappyFor(Lisa, Kate, e), 1)$. From this, (11.110), (11.111), (11.108), (11.118), $Lisa \neq Kate$, and E61 (which follows from (11.116)), we have

$$Happens(AddHappyFor(Lisa, Kate, WinLottery(Kate)), 1) \quad (11.120)$$

From (11.119), E9 (which follows from (11.113)), and DEC9, we have $HoldsAt(Joy(Kate, WinLottery(Kate)), 2)$. From (11.120), E13 (which follows from (11.113)), and DEC9, we have $HoldsAt(HappyFor\ (Lisa, Kate, WinLottery\ (Kate)), 2)$. ■

11.3 THE EPISTEMIC FUNCTIONAL EVENT CALCULUS

Jiefei Ma, Rob Miller, Leora Morgenstern, and Theodore Patkos introduced the *epistemic functional event calculus* (EFEC), an extension of the event calculus for reasoning about knowledge and action. Like the traditional event calculus, it supports reasoning about triggered events, concurrent events with cumulative and canceling effects, and nondeterministic effects of events. In addition, it supports reasoning about an agent's knowledge about past, present, and future fluent values and event occurrences.

The following shopping outlet scenario was used as a benchmark problem for the development of EFEC:

> *After a customer purchases an item at a computer terminal, the customer picks up the item at the correct merchandise counter, which is indicated on a display. Prizes are handed out at one of the counters after an item is successfully picked up. If the customer tries to collect the item from the wrong counter, the customer's purchase is canceled.*

EFEC allows us to represent that the customer's knowledge determines where the customer picks up the item. The customer acquires this knowledge through a *sensing action*.

Reasoning using EFEC can be performed using answer set programming. Rule sets for EFEC along with examples including the shopping outlet scenario are available for download (Ma, Miller, Morgenstern, & Patkos, 2013b).

EFEC is built using the *functional event calculus* (FEC), which extends the event calculus with non-truth-valued fluents. We first describe FEC, and then we proceed to a discussion of EFEC.

11.3.1 FEC BASICS

FEC uses the event, fluent, and timepoint sorts of the event calculus as well as a value sort, with variables v, v_1, v_2, ...

The predicates and functions of FEC are as follows:

$Happens(e, t)$: Event e happens or occurs at timepoint t.
$ValueOf(f, t) = v$: (Similar to $HoldsAt$.) The value of fluent f at timepoint t is v.
$CausesValue(e, f, v, t)$: (Similar to $Initiates$ and $Terminates$.) If event e occurs at timepoint t, then the value of fluent f is v after t.
$PossVal(f, v)$: The value v is a possible value of fluent f.

11.3.2 FEC AXIOMATIZATION

FEC consists of the following axioms and definitions.

Definition FEC1.

$$ValueCaused(f, v, t) \stackrel{\text{def}}{\equiv} \exists e\, (Happens(e, t) \wedge CausesValue(e, f, v, t))$$

Definition FEC2.

$$\begin{gathered} OtherValCausedBetween(f, v_1, t_1, t_2) \stackrel{\text{def}}{\equiv} \\ \exists t, v\, (ValueCaused(f, v, t) \wedge t_1 \le t < t_2 \wedge v \ne v_1) \end{gathered}$$

Axiom FEC3.

$$\begin{gathered} (ValueOf(f, t_1) = v \vee ValueCaused(f, v, t_1)) \wedge \\ t_1 < t_2 \wedge \neg OtherValCausedBetween(f, v, t_1, t_2) \Rightarrow \\ ValueOf(f, t_2) = v \end{gathered}$$

Axiom FEC4.

$$\begin{gathered} t_1 < t_2 \wedge OtherValCausedBetween(f, v, t_1, t_2) \wedge \\ \neg\exists t\, (t_1 \le t < t_2 \wedge ValueCaused(f, v, t)) \Rightarrow \\ ValueOf(f, t_2) \ne v \end{gathered}$$

Axiom FEC5.

$$ValueOf(f, t) = v \Rightarrow PossVal(f, v)$$

11.3.3 EFEC BASICS

EFEC defines knowledge using a *possible worlds* approach, in which an agent is defined to know a fact if and only if the fact is true in all worlds the agent considers possible. EFEC adds a collection of parallel timelines or *worlds* to the single timeline of the traditional event calculus. Each world is composed of a set of *instants*. In EFEC, the timepoint sort is no longer a subset of the real number sort, but rather an arbitrary sort.

EFEC adds the following to FEC:

- a world sort, with variables w, w_1, w_2, ...
- an instant sort, with variables i, i_1, i_2, ...
- a function $\langle w, i\rangle$, which represents the timepoint corresponding to world w and instant i
- a predicate $i_1 \prec i_2$, which represents that instant i_1 precedes instant i_2
- a predicate $t_1 < t_2$, which represents that timepoint t_1 precedes timepoint t_2
- an event $Sense(f)$, which represents that the agent senses the value of fluent f
- a constant W_a, which represents the actual world

EFEC also introduces a fluent $K(w)$, which represents the accessibility of world w. The expression $ValueOf(K(w_2), \langle w_1, i\rangle) = True$ represents that world w_2 is accessible from world w_1 at instant i.

11.3.4 EFEC AXIOMATIZATION

EFEC consists of the following axioms.

Axiom EFEC1.

$$\forall t\, \exists w, i\, (t = \langle w, i\rangle)$$

Axiom EFEC2.

$$\langle w_1, i_1\rangle = \langle w_2, i_2\rangle \Leftrightarrow w_1 = w_2 \wedge i_1 = i_2$$

Axiom EFEC3.

$$\langle w_1, i_1\rangle < \langle w_2, i_2\rangle \Leftrightarrow w_1 = w_2 \wedge i_1 \prec i_2$$

$\leq$, $>$, and $\geq$ are abbreviations defined appropriately in terms of $<$ and $=$.

Axiom EFEC4.

$$PossVal(K(w), v) \Leftrightarrow v = True \vee v = False$$

Axiom EFEC5.

$$ValueOf(K(w), \langle w, i\rangle) = True$$

Axiom EFEC6.

$$ValueOf(K(w_2), \langle w_1, i\rangle) = True \Rightarrow$$
$$ValueOf(K(w_1), \langle w_2, i\rangle) = True$$

Axiom EFEC7.

$$\begin{gathered} ValueOf(K(w_2), \langle w_1, i\rangle) = True \wedge \\ ValueOf(K(w_3), \langle w_2, i\rangle) = True \Rightarrow \\ ValueOf(K(w_3), \langle w_1, i\rangle) = True \end{gathered}$$

Axiom EFEC8. If the agent senses a fluent f in world w_1 at instant i and the value of f in w_1 at i is different from the value of f in a world w_2 at i, then the value of $K(w_2)$ is false in w_1 after i. That is, w_2 is not accessible from w_1 after i.

$$\begin{gathered} ValueOf(f, \langle w_1, i\rangle) \neq ValueOf(f, \langle w_2, i\rangle) \Rightarrow \\ CausesValue(Sense(f), K(w_2), False, \langle w_1, i\rangle) \end{gathered}$$

Axiom EFEC9. The agent knows in world w_1 at instant i_1 that the value of fluent f is not value v at instant i_2 if and only if, for every world w_2 accessible from w_1 at i_1, the value of f in w_2 at i_2 is not v.

$$\begin{gathered} KnowsValueIsNot(\langle w_1, i_1\rangle, f, i_2, v) \Leftrightarrow \\ \forall w_2\ (f \neq K(w_2) \wedge \\ (ValueOf(K(w_2), \langle w_1, i_1\rangle) = True \Rightarrow ValueOf(f, \langle w_2, i_2\rangle) \neq v)) \end{gathered}$$

Axiom EFEC10. The agent knows in world w at instant i_1 that the value of fluent f is value v_1 at instant i_2 if and only if, for every other possible value v_2 of f, the agent knows in w at i_1 that the value of f is not v_2 at i_2.

$$\begin{gathered} KnowsValueIs(\langle w, i_1\rangle, f, i_2, v_1) \Leftrightarrow \\ \forall v_2\ (PossVal(f, v_2) \wedge v_2 \neq v_1 \Rightarrow KnowsValueIsNot(\langle w, i_1\rangle, f, i_2, v_2)) \end{gathered}$$

Axiom EFEC11.

$$KnowsValue(t, f, i) \Leftrightarrow \exists v\ KnowsValueIs(t, f, i, v)$$

Axiom EFEC12. The agent knows in world w_1 at instant i_1 that event e happens at instant i_2 if and only if, for every world w_2 accessible from w_1 at i_1, event e happens in w_2 at i_2.

$$\begin{gathered} KnowsHappens(\langle w_1, i_1\rangle, e, i_2) \Leftrightarrow \\ \forall w_2\ (ValueOf(K(w_2), \langle w_1, i_1\rangle) = True \Rightarrow Happens(e, \langle w_2, i_2\rangle)) \end{gathered}$$

Axiom EFEC13. The agent knows in world w_1 at instant i_1 that event e does not happen at instant i_2 if and only if, for every world w_2 accessible from w_1 at i_1, event e does not happen in w_2 at i_2.

$$\begin{gathered} KnowsNotHappens(\langle w_1, i_1\rangle, e, i_2) \Leftrightarrow \\ \forall w_2\ (ValueOf(K(w_2), \langle w_1, i_1\rangle) = True \Rightarrow \neg Happens(e, \langle w_2, i_2\rangle)) \end{gathered}$$

Axiom EFEC14.

$$\begin{gathered} KnowsIfHappens(t, e, i) \Leftrightarrow \\ KnowsHappens(t, e, i) \vee KnowsNotHappens(t, e, i) \end{gathered}$$

The following axioms support representation of unconditional events, events conditional on knowledge, and triggered events.

Axiom EFEC15.

$$Perform(e, i) \Rightarrow Happens(e, \langle w, i \rangle)$$

Axiom EFEC16.

$$PerformIfValueKnownIs(e, f, v, i) \Rightarrow$$
$$(Happens(e, \langle w, i \rangle) \Leftrightarrow KnowsValueIs(\langle w, i \rangle, f, i, v))$$

Axiom EFEC17.

$$Triggered(e, t) \Rightarrow Happens(e, t)$$

Further axioms are required to ensure that sufficient accessible worlds exist to represent lack of knowledge in the initial instant and after nondeterministic event occurrences.

BIBLIOGRAPHIC NOTES

Beliefs and knowledge

Formal models of beliefs and knowledge are provided by R. C. Moore (1980, 1985a), Morgenstern (1988), E. Davis (1990, pp. 351-393), and Scherl and Levesque (2003). Possible worlds semantics is discussed by Hintikka (1962) and Kripke (1963). Fagin, Halpern, Moses, and Vardi (1995) provide a book-length treatment of reasoning about knowledge. Joseph Halpern, John McCarthy, and Heinrich Wansing debate the merits of modal logic versus first-order logic for representing mental states (J. Y. Halpern, 1999, 2000; McCarthy, 1997, 1999, 2000; Wansing, 1998). McCarthy (1998b) discusses giving robots knowledge of their own mental states. Elgot-Drapkin and Perlis (1990) and Elgot-Drapkin, Kraus, Miller, Nirkhe, and Perlis (1999) describe a logic called active logic for modeling the step-by-step reasoning processes of agents. Bratman (1987) treats the notions of beliefs, desires, and intentions. The epistemic functional event calculus (EFEC) was introduced by R. Miller, Morgenstern, and Patkos (2013) and Ma, Miller, Morgenstern, and Patkos (2013a).

Our axiomatization of beliefs in the event calculus is a simplified version of the one of Lévy and Quantz (1998). They develop a variant of the event calculus called the situated event calculus, which adds a situation argument to the event calculus predicates. They provide belief axioms such as

$$Initiates(adopt(p, s_f), believes(p, f, t_f), t_b, s_b) \leftarrow Holds\_at(f, t_f, s_f)$$

which represents that, if a person p adopts a situation s_f at timepoint t_b in situation s_b, and fluent f is true at timepoint t_f in situation s_f, then after t_b in s_b person p will believe that f is true at timepoint t_f.

Reactive behavior

Our representation of reactive behavior is similar to the stimulus-response rules of behaviorism (Watson, 1924). Brooks (1990) provides a language for reactive behaviors that compiles into code to be run on an architecture for controlling mobile robots (Brooks, 1985).

Goals and plans

Our event calculus axiomatization of goal-driven behavior is loosely based on agent control mechanisms such as those of Simon (1967); Fikes, Hart, and Nilsson (1972b); Wilensky (1983); Genesereth and Nilsson (1987, pp. 325-327); Bratman, Israel, and Pollack (1988); Mueller (1990); Rao and Georgeff (1992); Shoham (1993); Beaudoin (1994); Wooldridge (2000); and Winikoff, Padgham, Harland, and Thangarajah (2002). The hungry cat scenario is taken from Winikoff et al. (2002).

The persistence of plans and goals was pointed out by Lewin (1926/1951) and G. A. Miller, Galanter, and Pribram (1960). The study of planning is a large area within the field of artificial intelligence (Fikes & Nilsson, 1971; Ghallab, Nau, & Traverso, 2004; Newell & Simon, 1972; Sacerdoti, 1974, 1977; Wilensky, 1983) A collection of readings on planning is provided by Allen, Hendler, and Tate (1990). Logical formalizations of goals and plans as mental states are provided by Morgenstern (1988), Cohen and Levesque (1990), and E. Davis (1990, pp. 413-430).

Commonsense psychology

A number of researchers have analyzed the notions of the mental states and the events people use to model themselves and others. These notions are variously known as commonsense psychology, folk psychology, and theory of mind. Heider (1958) treats interpersonal relations and notions of perception, action, desire, pleasure, and sentiment. He provides relations such as "*p* SIMILAR TO *o* INDUCES *p* LIKES *o*, OR *p* TENDS TO LIKE A SIMILAR *o*" (p. 184).

Schank and Abelson (1977) provide a taxonomy of human goals and plans for achieving those goals. They classify goals according to whether they are to be repeatedly satisfied (S-HUNGER, S-SEX, S-SLEEP), are pursued for enjoyment (E-TRAVEL, E-ENTERTAINMENT, E-COMPETITION), are to be achieved (A-GOOD-JOB, A-POSSESSIONS, A-SOCIAL-RELATIONSHIPS), are states to be preserved (P-HEALTH, P-JOB, P-POSSESSIONS, P-SOCIAL-RELATIONSHIP), involve imminent threats (C-HEALTH, C-FIRE, C-STORM), or are instrumental to other goals (I-PREP, D-KNOW, D-PROX, D-CONT). They describe named plans and planboxes for achieving goals. For instance, a named plan USE(FOOD) for the goal of satisfying hunger (S-HUNGER) consists of knowing (D-KNOW) the location of some food, being near (D-PROX) the food, having physical control (D-CONT) over the food, and eating the food.

Smedslund (1997) provides a system of commonsense psychological propositions (45 definitions and 56 axioms). The propositions are broken up into the following categories: persons, acting, wanting, believing, feeling, interpersonal processes, intrapersonal processes, and personal change. Sample axioms are "*P wants X, if, and only if, other things equal, P prefers X to not-X*" (p. 29) and "*P is happy, if, and only if, and to the extent that, P believes that at least one of P's wants is being, or is going to be, fulfilled*" (p. 51). A. S. Gordon (2004) provides pre-formal representations of 372 strategies from 10 planning domains.

Emotions

Dyer (1983, pp. 103-139) presents a model of human emotions for story understanding. The representation of an emotion consists of a sign (positive or negative), story character experiencing the emotion, goal situation giving rise to the emotion, character toward whom the emotion is directed, strength, and expectation (expected or unexpected). A number of emotions can be represented using this scheme. Gratitude, for example, is represented as a positive emotion directed toward a character that caused a goal success for the character experiencing the emotion. Hope is represented as a positive emotion associated with an active goal that is expected to succeed.

Mueller (1990, pp. 53-56) extends Dyer's representation to model the influence of emotions on daydreaming and the influence of daydreaming on emotions. For example, an emotion of interest associated with an active goal initiates a stream of thought involving various scenarios for achieving the goal. If success of the goal is imagined, then an emotion of hope is initiated. Certain emotions are associated with certain personal goals. For example, embarrassment is defined as a negative emotion that results from a failed social esteem goal.

We formalize in the event calculus the model of emotions of Ortony, Clore, and Collins (1988). Our formalization is loosely based on the situation calculus formalization of O'Rorke and Ortony (1994). Some differences are as follows. O'Rorke and Ortony (p. 296) define emotional reactions using state constraints such as

$$joy(P, F, S) \leftarrow wants(P, F, S) \wedge holds(F, S)$$

rather than using trigger axioms and effect axioms. They represent that emotional reactions result directly from the fact that a fluent is true rather than from the fact that an agent believes that an event has occurred. They do not use real-valued functions for desirability, praiseworthiness, and anticipation, as suggested by Ortony, Clore, and Collins (1988, pp. 181-190).

Sanders (1989) presents a formalization of emotions using an extended version of Shoham's (1988) temporal logic. She treats emotions involving obligation, permission, and prohibition, namely, the emotions of anger, gratitude, approval, disapproval, shame, and guilt. For example, an axiom states that x is grateful to y if and only if (1) x believes that y performed an action a, (2) x wanted y to perform a, (3) x believes that a benefits x, and (4) x believes that y performed a without expecting anything in return. As in our axiomatization, the emotions of an agent arise from the beliefs of the agent. Unlike in our axiomatization, Sanders defines emotions using state constraints rather than trigger axioms and effect axioms.

EXERCISES

11.1 Formalize the reactive behavior of a driver. If a driver sees a squirrel in front of the car, then the driver will slam on the brakes.

11.2 If we remove the reactive behavior Axiom (11.16), then when the cat jumps onto the shelf the cat sticks to his original plan, jumping from the shelf to the table, ignoring the cookie on the shelf. Prove that if this axiom is removed, then the cat eats the food on the table.

11.3 Suppose that we remove Axiom (11.16), but we modify the axiomatization so that the agent drops plans that are nonoptimal. A plan is nonoptimal if a shorter plan exists. Provide appropriate modifications to Axioms A1–A11 so that nonoptimal plans are dropped. Note that when a nonoptimal plan is dropped, a new plan should be activated by Axiom A7. Modify other axioms as needed. Prove that the cat eats the food on the shelf.

11.4 Discuss the advantages and disadvantages of reactive rules versus rules that look for nonoptimal plans.

11.5 Extend the formalization in Section 11.2 so that (1) each emotion has an associated strength, (2) the strength is set appropriately based on desirability, praiseworthiness, and the likelihood of the anticipated event, (3) the strength of an emotion decays over time, and (4) an emotion is dropped when its strength drops below a certain threshold.

11.6 Formalize the operation of mental defense mechanisms (Freud, 1936/1946; Suppes & Warren, 1975). Use trigger axioms and determining fluents to apply transformations to the beliefs of an agent at each timepoint.

11.7 (Research problem) Extend the formalization in Section 11.1 to deal with multiple active goals and plans. Add priorities to active goals and a mechanism to select which active goal to work on at any moment.

11.8 (Research problem) McCarthy (1990) discusses a news story about a furniture salesman who is robbed. Formalize the knowledge necessary to represent and reason about this story.

11.9 (Research problem) Formalize processes of human daydreaming (Mueller, 1990) in the event calculus. Daydreaming consists of recalling past experiences, imagining alternative versions of past experiences, and imagining possible future experiences.