

Goals/Desires

:: straightforward – state of the art that needs to be accomplished, situations that are compatible with agent’s desires

$$(\text{Des } A \varphi)$$

:: desired accessibility relation is a **KD** modal system – the desired goals must not contradict

$$\Box\varphi \rightarrow \Diamond\varphi$$

$$\Box\varphi \rightarrow \neg\Diamond\neg\varphi$$

$$(\text{Des } A \varphi) \rightarrow \neg(\text{Des } \neg \varphi)$$

Intentions

:: an **intention** is something between the agents **state of mind** (belief) and the immediate action to be performed

:: unlike **desire/goal** an intention may be seen as agents immediate commitment to implementing an action

$$(\text{Int } A \varphi)$$

:: intention accessibility relation is a **KD** modal system – the intentions must not contradict

$$(\text{Int } A \varphi) \rightarrow \neg(\text{Int } \neg \varphi)$$

Intentions

:: properties of intentions:

- agents track the **success of their intentions**, and are inclined to try again if their attempts fail - intentions are persistent

$$(\text{Int } A \varphi) \rightsquigarrow \varphi$$

If an agents attempt to achieve fails, then all other things being equal, it will try an alternative plan to achieve

Intentions

:: properties of intentions:

- **satisfiability** – agents believe their intentions are possible; that is, they believe there is at least some way that the intentions could be brought about.

$$(\text{Int } A \varphi) \Rightarrow \text{EF}\varphi$$

- **intention-belief-inconsistency** – agents do not believe they will not bring about their intentions; it would be **irrational** of me to adopt an intention to if I believed was not possible.

$$(\text{Int } A \varphi) \wedge (\text{Bel } A \neg \text{EF}\varphi)$$

Intentions

:: properties of intentions:

- **intention-belief-incompleteness** - agent do not believe that their intention is possible to be achieved - may be rational behavior

$$(\text{Int } A \varphi) \wedge (\neg \text{Bel } A \text{EF}\varphi)$$

- agents admit that their intentions may not be implemented.

$$(\text{Int } A \varphi) \wedge (\text{Bel } A \text{EF}\neg\varphi)$$

- **time consistency** - $(\text{Int } A \varphi) \wedge (\text{Int } A \psi) \Rightarrow (\text{Int } A(\text{F}\varphi \wedge \text{F}\psi))$

Intentions

:: properties of intentions:

- **side-effects:** agents need not intend all the expected side effects of their intentions.

$$(\text{Bel } A \psi \Rightarrow \varphi) \wedge (\text{Int } A \psi) \wedge \neg(\text{Int } A \varphi)$$

is rational behaviour.

- intentions are not closed under implication - **side effect** problem.

I may believe that going to the dentist involves pain, and I may also intend to go to the dentist - but this does not imply that I intend to suffer pain!

Beleif-Desire-Intention Model

Framework for reasoning about formal abstract models of mental states (based on Theory of Practical Reasoning)

- :: contains representations (as objects, data structures, or whatever) of:
 - **beliefs**, which constitute its knowledge of the state of its environment (and perhaps also some internal state),
 - **desires**, which determine its motivation what it is trying to bring about, maintain, find out, etc.,
 - **intentions**, which capture its decisions about how to act in order to fulfill its desires (committed desires)

if $\varphi \in \mathcal{L}_{agent}$ then $\varphi, (\text{Bel } A \varphi), (\text{Des } A \varphi), (\text{Int } A \varphi) \in \mathcal{L}_{bdi}$

Properties of Rational Agents?

1. inevitables:

$$\begin{array}{ll}
 (\text{Int } A \text{ AG}\varphi) \Rightarrow (\text{Des } A \text{ AG}\varphi) & (\text{Des } A \text{ AG}\varphi) \Rightarrow (\text{Bel } A \text{ AG}\varphi) \\
 (\text{Des } A \text{ AG}\varphi) \Rightarrow (\text{Int } A \text{ AG}\varphi) & (\text{Int } A \text{ AG}\varphi) \Rightarrow (\text{Bel } A \text{ AG}\varphi) \\
 (\text{Bel } A \text{ AG}\varphi) \Rightarrow (\text{Des } A \text{ AG}\varphi) & (\text{Bel } A \text{ AG}\varphi) \Rightarrow (\text{Int } A \text{ AG}\varphi)
 \end{array}$$

2. options:

$$\begin{array}{ll}
 (\text{Int } A \text{ EF}\varphi) \Rightarrow (\text{Des } A \text{ EF}\varphi) & (\text{Des } A \text{ EF}\varphi) \Rightarrow (\text{Bel } A \text{ EF}\varphi) \\
 (\text{Des } A \text{ EF}\varphi) \Rightarrow (\text{Int } A \text{ EF}\varphi) & (\text{Int } A \text{ EF}\varphi) \Rightarrow (\text{Bel } A \text{ EF}\varphi) \\
 (\text{Bel } A \text{ EF}\varphi) \Rightarrow (\text{Des } A \text{ EF}\varphi) & (\text{Bel } A \text{ EF}\varphi) \Rightarrow (\text{Int } A \text{ EF}\varphi)
 \end{array}$$

Properties of Rational Agents

1. inevitables:

$$(\text{Int } A \text{ AG}\varphi) \Rightarrow (\text{Des } A \text{ AG}\varphi)$$

$$(\text{Des } A \text{ AG}\varphi) \Rightarrow (\text{Int } A \text{ AG}\varphi)$$

Properties of Rational Agents

1. inevitables:

$$(\text{Int } A \text{ AG}\varphi) \Rightarrow (\text{Des } A \text{ AG}\varphi)$$

$$(\text{Des } A \text{ AG}\varphi) \Rightarrow (\text{Int } A \text{ AG}\varphi)$$

2. options:

$$(\text{Int } A \text{ EF}\varphi) \Rightarrow (\text{Des } A \text{ EF}\varphi)$$

$$(\text{Des } A \text{ EF}\varphi) \Rightarrow (\text{Int } A \text{ EF}\varphi)$$

$$(\text{Des } A \text{ EF}\varphi) \Rightarrow (\text{Bel } A \text{ EF}\varphi)$$

$$(\text{Int } A \text{ EF}\varphi) \Rightarrow (\text{Bel } A \text{ EF}\varphi)$$