g-BDI: a graded intentional agent model for practical reasoning

–an application of the fuzzy modal approach to uncertainty reasoning–

Lluis Godo

Artificial Intelligence Research Institute (IIIA) - CSIC
Barcelona, Spain

Joint work with Ana Casali and Carles Sierra

Probability, Uncertainty and Rationality – Pontignano, November 1-3, 2009

Outline

- Introduction: BDI agent architectures and multi-context systems
- Background on the fuzzy logic approach to reasoning about uncertainty
- The g-BDI agent model
- A case study
- Concluding remarks

(Software) Agent theories and architectures

- **Theory:** a specification of an agent behaviour (properties it should satisfy)
  - The intentional stance (Dennet, 87)

    The behaviour can be predicted by ascribing certain mental attitudes e.g. beliefs, desires and rational acumen

- **Architecture:** software engineering model, middle point between specification and implementation (Wooldridge, 2001):
  - Logic-based: deliberative agents
  - Reactive: reactive agents
  - Layered: hybrid agents
  - Practical reasoning: BDI agents

    an explicitly representation of the agent's beliefs (B), desires (D) and intentions (I).

BDI agent models

The BDI agent model is based on M. Bratman's theory of *human practical reasoning* (reasoning to decide what and how to do), also referred to as Belief-Desire-Intention, or BDI:

• Intention and desire are both pro-attitudes (mental attitudes concerned with action), but intention is distinguished as a conduct-controlling pro-attitude: Intention = Desire + Commitment.

Several logical models to define and reason about BDI agents, e.g.

• Rao and Georgeff's BDI-CTL logic (1991) combines a multi-modal logic (with modalities representing beliefs, desires and intentions) with the temporal logic CTL*.

• Wooldridge (2000) has extended BDI-CTL to define LORA (the Logic Of Rational Agents), by incorporating an action logic, also allowing to reason about interaction in a multi-agent system.

g-BDI: a graded BDI agent model

Based on (Parsons et al., 98), we have proposed the g-BDI model that allows to specify agent architectures able to deal with the environment uncertainty and with graded mental (informational and proactive) attitudes.

- **Belief degrees** represent to what extent the agent believes a formula is true.
- **Degrees of positive or negative desires** allow the agent to set different ideal levels of preference or rejection respectively.
- **Intention degrees** also refer to preference but take into account the cost/benefit trade-off of reaching an agent's goal.

Working assumption

- Agents having different kinds of behavior can be modeled on the basis of the representation and interaction of these three attitudes.

Multi-Context Systems (Giunchiglia et al.)

MCSs exploits the idea of locality in reasoning and contain two basic components: contexts and bridge rules

A MCS is defined as $\left\langle \left\{ C_i \right\}_{i \in I}, \Delta_{br} \right\rangle$,

where

- Each context C_i is specified by
  - a logic $\langle L_i, A_i, \Delta_i \rangle$ where, L_i: language, A_i: axioms and Δ_i: inference rules
  - a theory $T_i \subseteq L_i$, encoding the available knowledge to C_i
- Δ_{br} is a set of bridge rules, i.e. rules of inference with premises and conclusions in different contexts

$$\frac{C_1 : \psi, C_2 : \varphi}{C_3 : \theta}$$

The deduction mechanism of a MCS is then based on the interplay between inter-context Δ_i and intra-context Δ_{br} deductions

g-BDI: a multi-context system based specification

A g-BDI agent is defined as a Multi-context System (MCS):

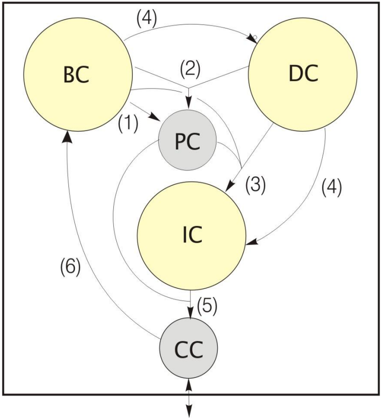
$$A_g = (\{BC, DC, IC, PC, CC\}, \Delta_{br})$$

where:

- The mental contexts represent: beliefs (BC), desires (DC) and intentions (IC).
- Two functional contexts are used for: Planning (PC) and Communication (CC).
- A suitable set of bridge rules (Δ_{br}) encode a particular pattern of interaction between Bs, Ds and Is

Such a MCS specification has advantages both from a logical and a software engineering perspectives (use of different logics, clear separation, modularity and efficiency, etc.)

g-BDI: a multi-context system based specification



Bridge Rule (5)

$$\frac{IC : (I_{\alpha_b}\varphi, i_{max}), PC : bestplan(\varphi, \alpha_b, P, A, c)}{CC : C(does(\alpha_b))}$$

g-BDI: graded logical framework

To represent and reason about the different graded mental attitudes in the g-BDI agent model, we use a fuzzy modal approach (Hájek et al.).

- *the belief / desire / intention degree of a Boolean proposition is considered as the truth-degree of a fuzzy (modal) proposition.*

- the algebraic semantics of different fuzzy logics can be used to characterize different models of measures.

This approach provides a uniform, quite powerful and flexible logical framework.

Outline

- Introduction: BDI agent architectures and multi-context systems
- Background on the fuzzy logic approach to reasoning about uncertainty
    - Fuzzy logic treatment of uncertainty
    - Probability logics
    - Possibilistic logics
- The g-BDI agent model
- A case study
- Concluding remarks

Graded representation of uncertainty

When belief is a matter of degree ...

B: set of events (Boolean algebra)

logical setting: $\mathbf{B} = \mathcal{L}/\equiv$
events as propositions (mod. logical equivalence)
$\top$ always true event,
$\bot$ always false event

Uncertainty, belief measures $\mu : \mathcal{L} \to [0, 1]$

$\mu(\varphi)$: quantifies an agent's confidence/belief on φ being true

(1) $\mu(\top) = 1, \mu(\bot) = 0$
(2) $\mu(\varphi) \leq \mu(\psi)$, if $\models \varphi \to \psi$
(3) $\mu(\varphi) = \mu(\psi)$, if $\models \varphi \equiv \psi$

Fuzzy measures (Sugeno) or Plausibility measures (Halpern)

Uncertainty measures: some classes of interest

(Finitely additive) Probability measures

Finite additivity: $P(\varphi \vee \psi) = P(\varphi) + P(\psi)$, whenever $\vdash \neg(\varphi \wedge \psi)$

- $P(\neg\varphi) = 1 - P(\varphi)$ (auto-dual)

Extension to conditional probabilities:

$P : \mathcal{L} \times \mathcal{L}^0 \to$ is a (coherent) conditional probability (De Finetti, Coletti and Scozzafava, ...):

(i) $P(\varphi \mid \varphi) = 1$, for all $\varphi \in \mathcal{L}^0$

(ii) $P(\cdot \mid \varphi)$ is a (finitely additive) probability for any $\varphi \in \mathcal{L}^0$

(iii) $P(\chi \wedge \psi \mid \varphi) = P(\chi \mid \varphi) \cdot P(\psi \mid \chi \wedge \varphi)$, for all $\psi \in \mathcal{L}$ and $\varphi, \chi \wedge \varphi \in \mathcal{L}^0$.

Uncertainty measures: some classes of interest

Possibility and Necessity measures

Possibility: $\Pi(\varphi \vee \psi) = \max(\Pi(\varphi), \Pi(\psi))$

Necessity: $N(\varphi \wedge \psi) = \min(N(\varphi), N(\psi))$

Dual pairs of measures (N, Π): when $\Pi(\varphi) = 1 - N(\neg\varphi)$

Representation in terms of possibility distributions

$$\pi : \Omega \to [0, 1]$$

$\pi(w) = 1$: w is totally plausible / preferred
$\pi(w) < \pi(w')$: w is less pausible / preferred than w'
$\pi(w) = 0$: w is impossible / rejected

$$N(\varphi) = \inf_{\omega \not\models \varphi} 1 - \pi(\omega) \qquad \Pi(\varphi) = \sup_{\omega \models \varphi} \pi(\omega)$$

Guaranteed posibility: $\Delta(\varphi) = \inf_{w \models \varphi} \pi(\varphi)$ min. level of satisfaction

Framing uncertainty reasoning in fuzzy modal theories

After P. Hájek (truth-degrees $\neq$ belief degrees!):

- for each crisp proposition φ, introduce a modality **P**

$$\mathbf{P}\varphi \text{ reads e.g. ``}\varphi\ \textbf{is probable}\text{''}$$

- $P\varphi$ is a gradual, fuzzy proposition: the higher is the probability of φ, the truer is $P\varphi$

- for φ a two-valued, crisp proposition one can define e.g.

$$\mathit{truth}{-}\mathit{value}(P\varphi) = \mathit{probability}(\varphi)$$

(which is different from $\mathit{truth}{-}\mathit{value}(\varphi) = \mathit{probability}(\varphi)$!!!)

Framing uncertainty reasoning in fuzzy modal theories

Crucial observation: laws and computations with probability (and many other measures) can be expressed by well-known fuzzy logic truth-functions on $[0, 1]$.

$$
\begin{aligned}
\mathit{Prob}(\varphi \vee \psi) &= \mathit{Prob}(\varphi) + \mathit{Prob}(\psi) - \mathit{Prob}(\varphi \wedge \psi) \\
&= \mathit{Prob}(\varphi) \oplus (\mathit{Prob}(\psi) \ominus \mathit{Prob}(\varphi \wedge \psi)) \\
\mathit{Prob}(\varphi \wedge \psi) &= \mathit{Prob}(\varphi) \cdot \mathit{Prob}(\psi \mid \varphi) \\
\mathit{Nec}(\varphi \wedge \psi) &= \min(\mathit{Nec}(\varphi), \mathit{Nec}(\psi)) \\
\mathit{Pos}(\varphi \vee \psi) &= \max(\mathit{Pos}(\varphi), \mathit{Pos}(\psi))
\end{aligned}
$$

Idea: axioms of different uncertainty measures on φ's to be encoded as *axioms of suitable fuzzy logic* theories over the $P\varphi$'s

Main systems of fuzzy logic

Extensions of Hájek's BL, whose standard semantics are given by the three outstanding t-norms:

Łukasiewicz logic: Ł $=$ BL $+$ $\neg\neg\varphi \equiv \varphi$

- $e(\varphi \&_Ł \psi) = \max(0, e(\varphi) + e(\psi) - 1)$
  $e(\varphi \to_Ł \psi) = \min(1, 1 - e(\varphi) + e(\psi))$

Gödel logic: G $=$ BL $+$ $\varphi \& \varphi \equiv \varphi$

- $e(\varphi \&_G \psi) = \min(e(\varphi), e(\psi))$
  $e(\varphi \to_G \psi) = 1$ if $e(\varphi) \leq e(\psi)$, $e(\varphi \to_G \psi) = e(\psi)$ otherwise

Product logic: $\Pi =$ BL $+$ $(\Pi 1), (\Pi 2)$

- $e(\varphi \&_\Pi \psi) = e(\varphi) \cdot e(\psi)$
  $e(\varphi \to_\Pi \psi) = \min(1, e(\psi)/e(\varphi))$

Łukasiewicz-Product logic: $Ł\Pi\frac{1}{2} = Ł + \Pi +$ few addional axioms

Definable connectives and truth functions

Connective	Definition	Truth function
$\neg_Ł\varphi$	$\varphi \to_Ł \overline{0}$	$1 - x$
$\varphi \oplus \psi$	$\neg_Ł\varphi \to_Ł \psi$	$\min(1, x + y)$
$\varphi \ominus \psi$	$\varphi \& \neg_Ł\psi$	$\max(0, x - y)$
$\varphi \equiv_Ł \psi$	$(\varphi \to_Ł \psi)\&(\psi \to_Ł \varphi)$	$1 - \lvert x - y \rvert$
$\varphi \wedge \psi$	$\varphi\&(\varphi \to_Ł \psi)$	$\min(x, y)$
$\varphi \vee \psi$	$(\varphi \to_Ł \psi) \to_Ł \psi$	$\max(x, y)$
$\Delta\varphi$	$\neg_\Pi\neg_Ł\varphi$	$\begin{cases} 1, & \text{if } x = 1 \\ 0, & \text{otherwise} \end{cases}$
$\neg_\Pi\varphi$	$\varphi \to_\Pi \overline{0}$	$\begin{cases} 1, & \text{if } x = 0 \\ 0, & \text{otherwise} \end{cases}$

A simple probability logic (HEG, 95), (Hájek, 98)

A two-level language:

(i) **Non-modal formulas:** φ, ψ, etc. , built from a set V of propositional variables $\{p_1, p_2, \ldots p_n, \ldots\}$ using the classical binary connectives $\wedge$ and $\neg$. The set of non-modal formulas will be denoted by $\mathcal{L}$.

(ii) **Modal formulas:** Φ, Ψ, etc. are built:
- from elementary modal formulas $P\varphi$, with $\varphi \in \mathcal{L}$
- using Lukasiewicz logic Ł connectives: $(\&_Ł, \rightarrow_Ł)$ and rational truth constants $\overline{r}$

Examples of FP- formulas: $\overline{0.8} \rightarrow_Ł P(\varphi \wedge \chi)$, $P(\neg\varphi) \rightarrow_Ł P(\chi)$,

Examples of non FP-formulas: $\varphi \rightarrow_Ł P\psi$, $\overline{0.5} \rightarrow_Ł P(P\varphi \wedge \chi)$

The logic $FP(CPC, RPL)$: axiomatization

- The set $Taut(\mathcal{L})$ of CPC tautologies
- Axioms of Rational Pavelka logic (Łukasiewicz logic + rational truth-constants) for modal formulas
- Probabilistic axioms:

(FP1)	$P(\varphi \to \psi) \to_{Ł} (P\varphi \to_{Ł} P\psi)$
(FP2)	$P(\varphi \vee \psi) \equiv (P\varphi \to_{Ł} P(\varphi \wedge \psi)) \to_{Ł} P\psi$
or equiv.	$P(\varphi \vee \psi) \equiv P\varphi \oplus (P\psi \ominus P(\varphi \wedge \psi))$
(FP3)	$P(\neg\varphi) \equiv \neg_{Ł} P(\varphi)$

- Deduction rules of FP(CPC, RPL) are *modus ponens* for $\to_{Ł}$ and
  - (-) *necessitation* for P: from φ derive $P\varphi$

The logic $FP(CPC, RPL)$: Semantics

Semantics: (weak) Probabilistic Kripke models $M = (W, e, \mu)$

- $e : W \times Var \to \{0, 1\}$
- $\mu : \mathcal{U} \subseteq 2^W \to [0, 1]$ probability s.t. the sets

  $$[\varphi] = \{w \in W \mid \|\varphi\|_{M,w}) = 1\} \text{ are } \mu\text{-measurable}$$

- atomic modal formulas: $\|P\varphi\|_M = \mu([\varphi])$
- compound modal formulas: $\|\Phi\|_{M,w}$ is computed from atomic using Łukasiewicz connectives

$M = (W, e, \mu)$ is a model of Φ if for any $w \in W$, $\|\Phi\|_{M,w} = 1$
(if Φ modal, it does not depend on w, only on μ)

Completeness of $FP(CPC, RPL)$: If T finite, $T \vdash_{FP} \Phi$ iff $T \models_{FP} \Phi$

Pavelka-style: $\sup\{r \mid T \vdash_{FP} \overline{r} \to_{Ł} \Phi\} = \inf\{\|\Phi\|_M \mid M \text{ model of } T\}$

$FP(CPC, RPL)$: a two-level framework

$$P\varphi \equiv_{Ł} \overline{0.3}, \quad P(\varphi \wedge \psi) \rightarrow_{Ł} P\chi\ , \quad \overline{0.6} \rightarrow_{Ł} P(\psi \vee \varphi), \ \ldots$$

uncertainty Łukasiewicz

---

events CPC

$$\neg(\psi \wedge \chi), \quad \varphi \wedge \psi \rightarrow \chi, \quad \varphi \vee (\psi \rightarrow \chi), \ \ldots$$

Generalization to other two-level logics

Fuzzy modal-like logics $FM(L_1, L_2)$

- L_1: logic of events, e.g. CPC, S5, Dynamic logic, Deontic logic, ...
- L_2: suitable fuzzy logic able to capture the corresponding intensional modality (probability, preference, belief, etc.)

Properties of $FM(L_1, L_2)$ obviously depend on those of L_1 and L_2

Some examples:

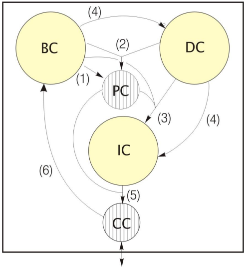
- conditional probability logic: $FP(CPC, \text{Ł}\Pi\frac{1}{2})$,
- belief function logic: $FP(S5, \text{Ł}\Pi\frac{1}{2})$
- possibilistic logics: $FN(CPC, G)$,
- graded deontic logics: $FP(SDL, RPL)$, $FN(SDL, G)$
- ...

Outline

- Introduction: BDI agent architectures and multi-context systems
- Background on the fuzzy logic approach to reasoning about uncertainty
- The g-BDI agent model
- A case study
- Concluding remarks

Outline

- Introduction: BDI agent architectures and multi-context systems
- Background on the fuzzy logic approach to reasoning about uncertainty
- The g-BDI agent model
    - Belief context
    - Desire Context
    - Itention Context
    - Bridge rules
    - Operational elements



- A case study
- Concluding remarks

The Belief Context

The purpose of this context is to model the agent's beliefs about the environment.

To represent knowledge related to action execution, we use Propositional Dynamic logic, PDL, (Fischer and Ladner, 79), as the base propositional logic

To account for the uncertainty on action execution, a probability-based approach and a necessity-based approach have been considered:

- $BC_{nec} = FN(PDL, G_\Delta(C))$ (strong standard completeness)
- $BC_{prob} = FP(PDL, RPL)$ (Pavelka-style completeness)

Typically, a theory will contain:

- quantitative formulas: $(B[\alpha]\varphi, 0.6)$ a shorthand for $\overline{0.6} \rightarrow_Ł B[\alpha]\varphi$ –the agent believes that the probability of φ being true after perfoming action α is at least 0.6–
- qualitative formulas: $B[\alpha]\varphi \rightarrow_Ł B[\beta]\varphi$

The Belief Context: BC_{prob}

Language:

- start from a PDL propositional language $\mathcal{L}_{PDL}$ built from a set of actions Π.
- introduce a belief operator B: if $\Phi \in \mathcal{L}_{PDL}$ then $B\Phi$ is a B-formula

  $P[\alpha]\varphi :=$ "φ is believed to be true after performing α"
- combine B-formulas with RPL connectives

Semantics: given by probabilistic Kripke structures:

$$\mathcal{M} = \langle W, \{R_\alpha : \alpha \in \Pi\}, e, \mu \rangle$$

where $\langle W, \{R_\alpha : \alpha \in \Pi\}, e \rangle$ is a regular Kripke model of PDL, and $\mu : F \to [0, 1]$ is a probability on a Boolean algebra $F \subseteq 2^W$

- $\|B\varphi\|_{\mathcal{M}} = \mu(\{w \in W \mid e(\varphi, w) = 1\})$

Axiomatics: PDL axioms + RPL axioms + (FP1), (FP2), (FP3)

Pavelka-style completeness

The Desire Context

The DC represents the agent's desires.

- Desires represent the *ideal agent's preferences*, regardless of the agent's current world and regardless of the cost involved in actually achieving them.
- Using the possibilistic approach to bipolar representation of preferences (Benferhat et al.) one can provide a fomal account of :
  - (graded) positive desires: what the agent would like to be the case.
  - (graded) negative desires: restrictions or rejections over the possible worlds it can reach.
  - indifference

The Desire Context

The Language $\mathcal{L}_{DC}$

- We start from a basic propositional language $\mathcal{L}$.
- To represent positive and negative desires over formulae of $\mathcal{L}$, we introduce two modal operators D^+ and D^-.

  $D^+\varphi :=$ "φ is positively desired"
  $D^-\varphi :=$ "φ is negatively desired " (or "φ is rejected").

- We use RPL to reason about modal formulas:
  - If $\varphi \in \mathcal{L}$ then $D^+\varphi, D^-\varphi \in \mathcal{L}_{DC}$
  - If $r \in \mathbb{Q} \cap [0,1]$ then $\bar{r} \in \mathcal{L}_{DC}$
  - If $\Phi, \Psi \in \mathcal{L}_{DC}$ then $\Phi \rightarrow_{Ł} \Psi \in \mathcal{L}_{DC}$ and $\neg_{Ł}\Phi \in \mathcal{L}_{DC}$

Notation: $(D^+\psi, r)$ will stand for $\bar{r} \rightarrow_{Ł} D^+\psi$

The Desire Context

Semantics - Intuition

- degrees of desire: a conservative approach

  minimum satisfaction / rejection levels

- the degree of (positive / negative) desire for a disjunction of goals $\varphi \vee \psi$ is taken to be the minimum of the degrees for φ and ψ.

  This is basically the characterizing property of the guaranteed possibility measures (Dubois-Prade et al.)

- the satisfaction degree of reaching both φ and ϕ can be strictly greater than reaching one of them separately. *The same for negative desires.*

The Desire Context

Semantics: intended models for $\mathcal{L}_{DC}$ are Kripke-like structures $\mathcal{M} = \langle W, e, \pi^+, \pi^- \rangle$, *Bipolar Desire models*, where:

- $\pi^+ : W \to [0,1]$ and $\pi^- : W \to [0,1]$ are positive and negative preference distributions over worlds.

$\pi^+(w) = 1$	full satisfaction	$\pi^-(w) = 1$	full rejection
$0 < \pi^+(w) < 1$	partial satisfaction	$0 < \pi^-(w) < 1$	partial rejection
$\pi^+(w) = 0$	indifference (nothing in favour)	$\pi^-(w) = 0$	indifference (nothing against)

- $\|D^+\varphi\|_{\mathcal{M}} = \inf\{\pi^+(w') \mid e(\varphi, w') = 1\}$

  $\|D^-\varphi\|_{\mathcal{M}} = \inf\{\pi^-(w') \mid e(\varphi, w') = 1\}$

- e is extended to compound modal formulae by means of the usual truth-functions for Łukasiewicz connectives.

$\mathcal{M} \models \Phi$ and $T \models_{\mathcal{M}} \Phi$ as usual.

The Desire Context

We define the *Basic logic for DC* (DC logic) as follows:

Axioms and rules:

(CPC) Axioms and rules of classical logic for non-modal formulas

(RPL) Axioms and rules of Rational Pavelka logic for modal formulas

(DC$^+$) $\quad D^+(\varphi \vee \psi) \equiv_{Ł} D^+\varphi \wedge_{Ł} D^+\psi$

(DC$^-$) $\quad D^-(\varphi \vee \psi) \equiv_{Ł} D^-\varphi \wedge_{Ł} D^-\psi$

Introduction of D^+ and D^- for implications:

(ID$^+$) from $\varphi \to \psi$ derive $D^+\psi \to_{Ł} D^+\varphi$

(ID$^-$) from $\varphi \to \psi$ derive $D^-\psi \to_{Ł} D^-\varphi$.

Soundness and Completeness

The above axiomatization is correct with respect to the defined semantics and is complete as well for finite theories of modal formulas.

The Desire Context

Encoding different types of desires in a DC theory: some examples

Suppose María looks for possible touristic destinations matching her preferences:

- She likes beach and mountain destinations, beach is a bit more preferred
  $(D^+ beach, 0.8)$, $(D^+ mountain, 0.7)$
  $D^+ mountain \rightarrow_Ł D^+ beach$
- She is totally indifferent to destinations with or without a zoo
  $\neg_Ł D^+ zoo, \neg_Ł D^+ \neg zoo$
  $\neg_Ł D^- zoo, \neg_Ł D^- \neg zoo$
- She does not want to travel with bus.
  $(D^- bus, 0.9)$
- She likes the train but she is a bit afraid of possible delays. Anyway, she does not discard the train, but the plane would be preferable.
  $(D^+ train, 0.5)$, $(D^+ \neg train, 0.2)$, $\neg_Ł D^- train$
  $D^+ train \rightarrow_Ł D^+ plane$

The Desire Context

Too much freedom?

For some classes of problems we may want to restrict the allowed assessments of degrees of positive and negative desires.

- Different axiomatic extensions can be proposed to show how different consistency constraints can be added to the basic DC logic, both at the semantical and syntactical levels, while preserving completeness.
- One think of such extensions as modelling different types of agents.

DC_1 Schema

It may be natural in some domain applications to forbid to simultaneously positively (negatively) desire φ and $\neg\varphi$

These constraints amount to require in the intended models:

- $\min(\|D^+\varphi\|_M, \|D^+\neg\varphi\|_M) = 0$, and
- $\min(\|D^-\varphi\|_M, \|D^-\neg\varphi\|_M) = 0$

At the level of Kripke structures, this corresponds to:

- $\inf_{w\in W} \pi^+(w) = 0$, and
- $\inf_{w\in W} \pi^-(w) = 0$

At the syntactic level, this is captured by :

$$(\mathrm{DC1}^+) \quad \neg_{Ł}(D^+\varphi \wedge_{Ł} D^+(\neg\varphi))$$
$$(\mathrm{DC1}^-) \quad \neg_{Ł}(D^-\varphi \wedge_{Ł} D^-(\neg\varphi))$$

DC_2 Schema

The logical schema DC_1 does not put any restriction on positive and negative desires for a same goal.

Benferhat et al.'s coherence condition: An agent cannot desire to be in a world more than the level at which it is tolerated, i.e. not rejected.

Translated to our framework, it amounts to require:

- $\forall w \in W,\ \pi^+(w) \leq 1 - \pi^-(w)$

This captured at the syntactical level by:

$$(\mathrm{DC}_2) \quad \neg_{Ł}(D^+\varphi \otimes D^-\varphi)$$

where $\otimes$ is Łukasiewicz strong conjunction.

DC_3 Schema

An stronger consistency condition between positive and negative preferences may be considered:

If an agent rejects (desires) to be in a world to some extent, it cannot be positively desired (rejected) at all

At the semantical level, this amounts to require:

$$\min(\pi^+(w), \pi^-(w)) = 0$$

At the syntactic level:

(DC3) $(D^+\varphi \ \wedge_Ł D^-\varphi) \rightarrow_L \bar{0}$

The Intention Context
From Desires to Intentions

In the g-BDI agent model, positive and negative desires are used as pro-active and restrictive elements respectively in order to set up intentions.

- intentions cannot depend just on the satisfaction of reaching a goal φ (represented by $D^+\varphi$) but also on the state of the world and the cost of transforming it into a world where the formula φ is true.

- a graded representation allows us to define the strength of an intention as a measure of the cost/benefit relationship of the feasible actions the agent can take toward the intended goal.

Intention context

Language $\mathcal{L}_{IC}$:

- Elementary modal formulae $I_\alpha \varphi$, where $\alpha \in \Pi^0 \subset \Pi$ (finite)

  The truth-degree of $I_\alpha \varphi$ will represent the strength the agent intends φ by means of the execution of the particular action α.

  - intended semantics: trade-off between preference and cost –

- $\mathcal{L}_{IC}$ formulas are built from $Var_{cost} = \{c_\alpha\}_{\alpha \in \Pi^0}$ and elementary modal formulaes $I_\alpha \varphi$, using Rational Łukasiewicz logic (Gerla, 01) connectives .

Intention context

Semantics: intended models will be enlarged Kripke structures $M = \langle W, e, \pi^+, \pi^-, \{\pi_\alpha\}_{\alpha \in \Pi^0} \rangle$ where $\pi_\alpha : W \times W \to [0,1]$ is a *utility* distribution corresponding to α

> $\pi_\alpha(w, w')$*: utility degree of applying* α *to transform world* w *into world* w'*.*

- $e(w, I_\alpha \varphi) = \inf\{\pi_\alpha(w, w') \mid w' \in W, e(w', \varphi) = 1\}$

Additional axioms and rules

1. (DC) axiom for I_α modalities: $I_\alpha(\varphi \vee \psi) \equiv_{Ł} I_\alpha \varphi \wedge_{Ł} I_\alpha \psi$
2. introduction of I_α for implications: from $\varphi \to \psi$ derive $I_\alpha \psi \to_{Ł} I_\alpha \varphi$ for each $\alpha \in \Pi^0$

Theorem

Let T *be a finite theory of modal formulas and* Φ *a modal formuls. Then* $T \vdash_{IC} \Phi$ *iff* $T \models_{\mathcal{M}_{IC}} \Phi$*.*

Intention context

A particular semantics

Intended semantics: $e(w, I_\alpha\varphi) =$ trade-off between the degree of ideal preference (positive desire) of φ and $1-$ the cost of achieving φ by performing α at w.

Example: assume we take Rational Łukasiewicz logic (Gerla) to reason about modal formulas (Ł logic + δ_n's). Then consider the additional axiom:

$$I_\alpha\varphi \equiv_{Ł} \delta_2 D^+\varphi \oplus \delta_2 c_\alpha$$

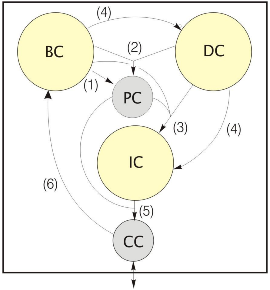
This axiom is valid in extended structures $M = \langle W, e, \pi^+, \{\pi_\alpha\}_{\alpha\in\Pi^0}\rangle$ iff

$$\pi_\alpha(w, w') = \pi^+(w') + 1 - e(w, c_\alpha)$$

. . . but bridge rules offer a lot of flexibility . . .

Bridge Rules

Bridge rules allow to embed results from a theory into another, they are part of the deduction mechanism of the g-BDI agent.



Intention generation rule:

$$BR(3) \quad \frac{DC : (D^{+}\varphi, d), \quad BC : (B[\alpha]\varphi, r), \quad PC : fplan(\varphi, \alpha, P, A, c)}{IC : (I_{\alpha}\varphi, f(d, r, c))}$$

$$f(d, r, c) = r \cdot (w_d d + w_c(1 - c))$$

Bridge Rules

Other bridge rules that can be used:

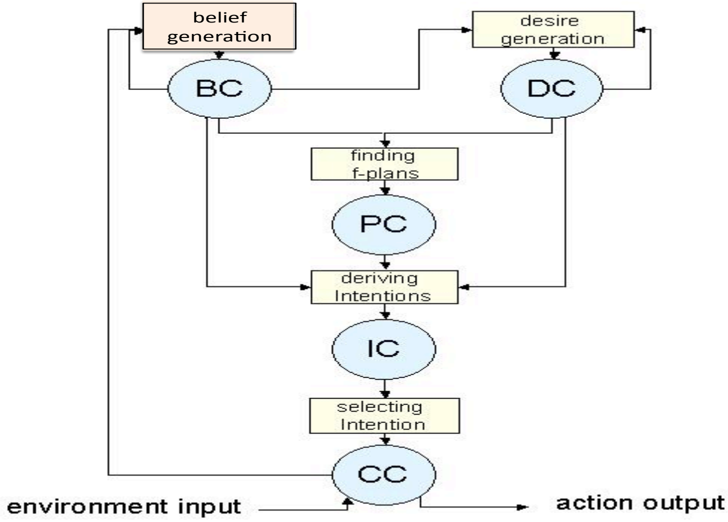
- Bridge rules to represent realism relations between mental attitudes (Cohen and Levesque), e.g.

$$\frac{BC : \neg B\varphi}{DC : \neg D^{+}\varphi} \qquad \frac{DC : \neg D^{+}\varphi}{IC : \neg I\varphi}$$

- Bridge rules to generate desires in a dynamic way: Rahwan and Amgoud's *Desire-Generation Rules*

$$\frac{BC : (B\varphi_1 \wedge ... \wedge B\varphi_n, b), \quad DC : (D^{+}\psi_1 \wedge ... \wedge D^{+}\psi_m, c)}{DC : (D^{+}\psi, d)}$$

How does the g-BDI model work?



Example

María, a tourist, activates a personal agent based on the g-BDI agent model, to get a tourist package that satisfies her preferences. She would be very happy going to a mountain place (m), and rather happy practicing rafting (r). On top of this, she wouldn't like to go farther than 1000km from Buenos Aires (f) where she lives.

The recommender agent takes all desires expressed by María and follows the steps:

- **Desire generation:** the user interface that helps her express these desires ends up generating a desire theory for the DC as follows:

$$\mathcal{T}_{\mathcal{D}} = \{(D^+m, 0.9), (D^+r, 0.6), (D^+(m \wedge r), 0.96), (D^-f, 0.7)\}$$

Example

- **Beliefs generation:** with the tourism plans offered, the tourism domain and its beliefs about how these packages can satisfy the user's preferences ($\mathcal{T}_{\mathcal{B}}$).

  - Plans: *Mendoza* (*Me*), *SanRafael* (*Sr*), *Cumbrecita* (*Cu*), ...

  - Costs: $c(Me) = 0.60$, $c(Sr) = 0.70$, $c(Cu) = 0.55$, ...

  - Beliefs: $(B[Me]m, 0.7)$, $(B[Me]r, 0.6)$, $(B[Me]m \wedge r, 0.6)$,
    $(B[Sr]m, 0.5)$, $(B[Sr]r, 0.6)$, $(B[Sr]m \wedge r, 0.5)$, ...

- **Looking for feasible packages:** from this set of positive and negative desires ($\mathcal{T}_{\mathcal{D}}$) and domain knowledge ($\mathcal{T}_{\mathcal{B}}$) the PC looks for *feasible plans*, that are believed to achieve positive desires (m, r, $m \wedge r$) but avoiding the negative desire (f) as a post-condition.

  - *Mendoza* (*Me*) and *SanRafael* (*Sr*) are feasible plans for the combined goal $m \wedge r$, while *Cumbrecita* (*Cu*) is feasible only for m.

Example

- **Deriving the Intention formulae:** the intention degrees for satisfying each desire m, r and $m \wedge r$ by the different feasible plans are computed by the bridge rule that trades off the cost and benefit of satisfying a desire by following a plan. The IC context is filled up with the following formulas:

$$
\begin{array}{lll}
\mathcal{T}_{IC} = \{ & (I_{Me}(m \wedge r), 0.675), & (I_{Sr}(m \wedge r), 0.625), \\
& (I_{Me}(m), 0.60), & (I_{Me}(r), 0.50), \\
& (I_{Sr}(m), 0.55), & (I_{Sr}(r), 0.45), \\
& (I_{Cu}(m), 0.625) \;\; \} &
\end{array}
$$

- **Selecting Intention-plan:** the agent decides to recommend the plan *Mendoza* (Me) since it brings the best cost/benefit relation (represented by the intention degree 0.675) to achieve $m \wedge r$, satisfying also the tourist's constraints.

Outline

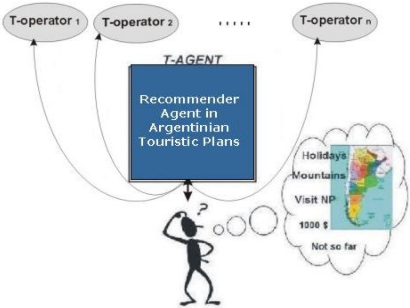
- Introduction: BDI agent architectures and multi-context systems
- Background on the fuzzy logic approach to reasoning about uncertainty
- The g-BDI agent model
- A case study
- Concluding remarks

A Case Study

A Tourism Recommender System

Goal: development (as a proof-of-concept) of a tourism recommender system to recommend the best tourist packages on Argentinian destinations according to different user's preferences and restrictions, provided by different tourist operators.

Main task: design of a Travel Assistent agent (*T-agent*), implementation, validation and experimentation



T-Agent implementation

Communication context: is the *T-Agent* interface, interacting with

- P-Agents (Tourist Operators), updating the information about current packages
- the user (tourist customer) that is looking for recommendation (Web service application).

RECOMENDADOR DE TURISMO

USUARIO			
NOMBRE	MARIA		
PREFERENCIAS			
☑ ZONA	PATAGONIA	9	
☐ RECURSOS NATURALES	MAR	5	
☐ INFRAESTRUCTURA	MUSEO ARQUEOLOGIA	5	
☑ TRANSPORTE	AVION	7	
☑ ALOJAMIENTO	APART	6	
☐ ACTIVIDADES	CABALGATA	5	
FRECUENCIA DE LA ACTIVIDAD	BAJA		
RESTRICCIONES			
☑ COSTO	2000		
☐ DISTANCIA A RECORRER	0		
☐ DIAS	0		
TIPO DE RESTRICCIONES	FLEXIBLE		
PARAMETROS DE LA CONSULTA			
PRIORIDAD	SATISFACCION DE RESTRICCIONES		

Enviar

Experimentation and validation

Made over 52 queries, 35 different users (students)

Some conclusions (to be taken cautiously):

1. the g-BDI model has been proved useful to build concrete agents in real world applications.
2. the *T-Agent* recommended rankings (over 40 Tourism packages) are in most of the cases close to the user's own rankings.
3. g-BDI agent architecture allows us to engineer agents having different behaviours by suitably tuning some of its components.
4. the distinctive feature of recommender systems modelled using g-BDI agents, which is using graded mental attitudes, allows them to provide better results than those obtained with non-graded BDI models.

Concluding remarks

Future Work

- **Social aspects:**
  An important topic for further work is to consider how to evaluate the trust-reputation in other agents, and how the agent updates this model along time.
- **Dynamic aspects:**
  To model agents that interact in dynamic environments, the g-BDI agent should be extended to account for a temporal dimension in what regards her beliefs, desires and intentions.
- **Revision in g-BDI Agents:**
  - g-BDI agents must be able to deal with contextual inconsistencies (revision mechanism, argumentation system?)
  - need of a general process for multi-context system revision and then specialize it for the g-BDI agent model.

Thank you !