Argumentation Basics

- An argument is:
  - A set of sentences/rules, S, in some background logic $(L, \vdash)$: from which we can derive a conclusion (I.e. $S \vdash \phi$)
- Attacking Relation:
  - Specifies when one argument (i.e. a set S_1 of rules) attacks another argument S_2 e.g. when:
    - they have some contrary conclusion and S_1 is "as strong" as S_2.
- Admissibility/Acceptability criterion:
  - Selects appropriate arguments (from a given corpus), called admissible/acceptable, that "behave well" under their attacks
- An argument S is Admissible:
  - S is conflict free (i.e. it does not attack itself) and
  - S attacks (counter-attacks) all its attacks
- Credulous or Skeptical Reasoning:
  - A conclusion holds in one or all admissible/acceptable extensions

Argumentation Realization

STEP 1

- **Preference based argumentation**
  - The attacking relation is defined in terms of a priority structure on the arguments.

Preference Based Argumentation (1)

- What is an attack on S_1?
  - An argument S_2 with **contrary** claims (either for the original top-level decision or for the supporting ones)
  - Where S_2 is also **NOT less preferable**.
- What is "**less preferable**"?
  - Contains weaker components (links)
- What is a "**weaker** component/link"?

Preference Based Argumentation (2)

- What is a "**weaker** component/link"?
  - This is stated explicitly in the theory/knowledge, eg.
    - "Social responsibility is stronger than personal gain"
    - "Later laws are stronger than earlier ones"
    - "Later events have stronger information than earlier ones"
    - "Specific information is stronger than general information"
  - **LOCALLY specified and lifts via the argumentation to give GLOBAL (overall preferred) decisions.**

=> MODULARITY of Representation

=> MODULARITY of Design and Architecture of Agents

Preference Based Argumentation (3)

- What is a “weaker component/link”?
  - This weaker/stronger notion is not fixed but conditional e.g.:
    - “A law is stronger than the another WHEN this is passed later”
    - “Accepting a requested task is stronger than carrying out your own task WHEN the request comes from a superior”
  - This dynamic nature of preferences/attacking is vital in a changing environment
    - **Adaptability of argumentative reasoning**
    - **Personalization of argumentative reasoning**

Argumentation Realization
STEP 2

- Realizations in a logical framework
- An argument is a set of sentences to support a conclusion in some background monotonic logic $(\mathcal{L}, \vdash)$:
  - AF = <T, Att>, where T is a theory in some logic
  - Given a subset of sentences S we can derive conclusions ($S \vdash \phi$)
    - **These conclusions are the positions of the argument**

The Attacking Relation

- An attacking relation is realized between sets of sentences, ϕ and ψ, as:

  1) ϕ and ψ have a contrary conclusion

  2) Strength Relation via Priorities:
  - $\phi \preceq_{DYN} \psi$ iff $(\exists r \in \phi, r' \in \psi : \phi \vdash r' < r) \Rightarrow (\exists r \in \phi, r' \in \psi : \psi \vdash r < r')$

  Then Att(ψ, ϕ), i.e. ψ attacks ϕ.

- Strong and Weak attacks.

Logic Programming without Negation as Failure (LPwNF)

- LPwNF:
  - A concrete scheme of the abstract argumentation framework which uses explicit negation for conflict.
  - Labelled rules of the form Label: $L \leftarrow L_1, \ldots, L_n$ where $L, L_1, \ldots, L_n$ are positive or explicit negative literals and Label is a functional term.
- Extensions:
  - Generalized the attacking relation to be dynamic.
  - Integrated abduction.

Logic Programming without Negation as Failure (LPwNF)

- Horn background logic:
  - Rules: $L \leftarrow L_1, \ldots, L_n$ where $L, L_1, \ldots, L_n$ literals $L_i = (\neg)A_i$
  - Contrary given by classical negation $\neg$
  - Priority relation ">" on rules of the theory

Example

p← q, not r "p holds if q holds unless r holds"

R_1: p← q
R_2: ¬p←r
$R_2 > R_1$

- Attacking relation given by:
  - S attacks S' iff there exist L and $S_1 \subseteq S$, $S'_1 \subseteq S'$ s.t.:
    - $B \cup S_1 \vdash_{min} L$ and $B \cap S'_1 \vdash_{min} \neg L$
    - $S_1 \supseteq S'_1$ (If S_1 has a rule of lower priority then it also has one of higher priority)

An Example of Argumentation Theory Policy

- Decision policy of a seller agent

r1: sell(Prd, Ag, high-price) ← pay-card(Ag, Prd)
r2: sell(Prd, Ag, high-price) ← pay-install(Ag, Prd)
r3: sell(Prd, Ag, low-price) ← pay-cash(Ag, Prd)
r4: ¬sell(Prd, Ag, P2) ← sell(Prd, Ag, P1), P2≠P1

  - Priority: r1 > r2, r1 > r3 , r2 > r3

Argumentation with Roles and Context

- **Default Context**↔definition of roles
  - Market: normal, regular customer
- **Specific Context**
  - High season, sales season
- **Example Agent theory: T=($\mathcal{T}$, $\mathcal{P}_R$, $\mathcal{P}_C$)**

  R1: h-p(r1(Prd, Ag), r3(Prd, Ag))

  R2: h-p(r3(Prd, Ag), r1(Prd, Ag)) ← regular(Ag), buy_2(Ag, Prd)

  R3: h-p(r3(Prd, Ag), r1(Prd, Ag)) ← regular(Ag), late_del(Ag, Prd)

  C1: h-p(R1(Prd, Ag), R2(Prd, Ag)) ← high-season

  C3: h-p(R2(Prd, Ag), R3(Prd, Ag)) ← special-product(Prd)

- MODULARITY of representation

Personality Theory

- Maslow's (default) Hierarchy of Human Needs (**"other things being equal"**)
  - Physiological
  - Safety
  - Affiliation or Social
  - Achievement or Ego
  - Self-Actualization or Learning
- Argumentation based preference policy for goal decision

Agent Deliberation on Needs and Motivations (2)

- **Satisfied and Critical Needs**
  - S_j: the set of conditions, evaluated in the theory T of the agent, under which the agent considers that his needs pertaining to motivation m_j are **satisfied**
  - N_j: the set of conditions, evaluated in the theory T of the agent, under which the agent considers that his needs pertaining to motivation m_j are **critical**
  - S_j and N_j are disjoint

- **Default motivation preference theory of Agent**
  - R^1_{ij}: h-p(G_i, G_j) ← N_i
  - R^2_{ij}: h-p(G_i, G_j) ← ¬S_i, ¬N_j

  where G_i and G_j are any two potential goals ($i \neq j$) of the agent associated to motivations m_i and m_j respectively

Agent Deliberation on Needs and Motivations (3)

- An agent theory expressing his profile on needs is a theory $T=(\mathcal{T}, \mathcal{P}_M, \mathcal{P}_C)$ where:
  - $\mathcal{T}$ is a background theory to compute S_j and N_j
  - $\mathcal{P}_M$ contains the rules:
    - R^1_{ij}: h-p(G_i, G_j) ← N_i
    - R^2_{ij}: h-p(G_i, G_j) ← ¬S_i, ¬N_j
  - For each pair of rules R^k_{ij}, R^k_{ji} in $\mathcal{P}_M$ we have the following rules in $\mathcal{P}_C$:
    - H^k_{ij}: h-p(R^k_{ij}, R^k_{ji})←true
    - E^k_{ji}: h-p(R^k_{ji}, R^k_{ij}) ←sc^k_{ij}
    - C^k_{ij}: h-p(E^k_{ij}, H^k_{ji})←true

    where sc^k_{ij} are special conditions whose truth can be evaluated in $\mathcal{T}$.
  - The rules H^k_{ij} are called the **basic hierarchy** of the theory T and the rules E^k_{ji} the **exception policy** of the theory T.

Agent Deliberation on Needs and Motivations (4)

- These profiles $T=(\mathcal{I}, \mathcal{P}_{\mathcal{M}}, \mathcal{P}_{C})$ capture via $\mathcal{P}_{C}$ different personalities:
  - Selfish, Altruist, etc
  - But sensitive to special circumstances where the default behaviour is over-written

Example

- G_1=fill up (N_1), G_2 =help in work ($\neg S_3$) $\rightarrow$ G_1
- G_1=fill up ($\neg S_1$), G_2 =help in work ($\neg S_3$) $\rightarrow$ G_1
  - Dilemma $\rightarrow$ G_1 or G_2 according to the basic profile of the agent
- G_1=fill up (N_1), G_2=help injured (N_3) $\rightarrow$ G_1
  - Dilemma $\rightarrow$ G_1 or G_2 according …
    - If injured=child (special condition: case 31) $\rightarrow$ G_2

Capabilities and Personality(1)

- The Personality can influence the decision making of the agent associated to his different capabilities

- Example: Decide within the problem solving module which requested task to perform according to his "professional" policy and his personality

➢ Professional Policy

r1(A, T1, A1): perform(A, T1, A1) ← ask(A1, T1, A)
r2 (A, T1, T2, A1): ¬perform(A, T1, A1)← perform(A, T2, self)
R1: h-p(r1(A, T1, A1), r2(A, T1, T2, A1)) ←higher_rank(A1, A)
R2: h-p(r2(A, T1, T2, A1), r1(A, T1, A1)) ← competitor(A1, A)
C1: h-p(R1(A, T1, T2, A1), R2(A, T1, T2, A1)) ← common _project(A, T1, A1)
C2: h-p(R2(A, T1, T2, A1), R1(A, T1, T2, A1)) ← urgent(A, T2)

➢ Personality Policy: The case of a selfish agent

R^2_{43}: h-p(G_4, G_3) ← ¬S_4, ¬N_3
R^2_{34}: h-p(G_3, G_4) ← ¬S_3, ¬N_4
H^2_{43}: h-p(R^2_{43}, R^2_{34})←true
E^2_{34}: h-p(R^2_{34}, R^2_{43}) ←dangerous_for_company(G_4)
C^2_{34}: h-p(E^2_{34}, H^2_{43})←true

Capabilities and Personality(2)

- "Professional" and personality policies can be in conflict
- The method of conflict resolution exploits the agent's ability to synthesize argumentation and abductive reasoning
- It assumes that these conflicts occur due to lack of information
- Given two opposing goals G_1 and G_2 there are three possible cases of such a conflict
  - Case 1: G_1 and G_2 are skeptical conclusions of the professional theory of a module and the personality theory respectively
  - Case 2: G_1 is a skeptical conclusion of the professional theory of a module and G_2 is a credulous conclusion (and hence so is G_1) of the personality theory (or vice-versa)
  - Case 3: G_1 and G_2 are credulous conclusions of the professional theory of a module and the personality theory respectively

Capabilities and Personality(3)

- Mechanism for resolving conflicts
  - Suspend decision
  - Deliberate on goals to find supporting information that would strengthen or weaken the conclusions of the separate theories
  - Evaluate if possible (some of) this supporting information in the external environment, and if this results in:
    - Case 1: then the agent chooses one of the goals according to a simple preference for or against the personality choice (i.e. given by the designer)
    - Case 2: then the agent decides for the goal that is skeptically true
    - Case3: then the agent selects randomly one of the two goals