http://www.diva-portal.org

Postprint

This is the accepted version of a paper presented at *Pre-AAMAS2018 Workshop on Formal Methods and Logical Aspects of Multi-Agent Systems (FMLAMAS 2018), 10 July 2018, Stockholm, Sweden*.

Citation for the original published paper:

Morveli Espinoza, M., Nieves, J C., Possebom, A T., Puyol-Gruart, J., Tacla, C A. (2018)
Resolving Incompatibilities between Procedural Goals: An Argumentation-based Approach
In: Riccardo De Masellis, Valentin Goranko (ed.),

N.B. When citing this work, cite the original published paper.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-152009

Resolving Incompatibilities between Procedural Goals: An Argumentation-based Approach*

Mariela Morveli Espinoza
UTFPR, Curitiba, Brazil
morveli.espinoza@gmail.com

Juan Carlos Nieves
Umea University, Umea, Sweden
jcnieves@cs.umu.se

Ayslan T. Possebom
UTFPR, Curitiba, Brazil
possebom@gmail.com

Josep Puyol-Gruart
IIIA-CSIC, Barcelona, Spain
puyol@iiia.csic.es

Cesar Augusto Tacla
UTFPR, Curitiba, Brazil
cesar.tacla@gmail.com

ABSTRACT

During the deliberation phase in practical reasoning, an intelligent agent generates a set of pursuable goals (or desires) and then selects which of them he commits to achieve (intentions). When several pursuable goals are generated, it may cause that some incompatibilities arise between them. In this work, we focus on the definition, identification and resolution of these incompatibilities. The suggested approach considers the three forms of incompatibility introduced by Castelfranchi and Paglieri [1], namely the terminal incompatibility, the resources incompatibility and the superfluity. We characterise computationally these forms of incompatibility by means of arguments that represent the plans that allow an agent to achieve his goals. The incompatibility between goals is defined based on the conflicts between their plans, which are represented by means of attacks in an argumentation framework. For the goals selection, we propose to use abstract argumentation theory by applying argumentation semantics.

1. INTRODUCTION

This work concerns with he process of goals selection during practical reasoning; i.e. on deciding which consistent set of goals the agent will commit to. This process includes identifying incompatibilities between pursuable goals and resolving such incompatibilities. We specifically work with procedural goals and resource-bounded agents.

Given that an intelligent agent may generate multiple pursuable goals, some incompatibilities between these goals could arise. Since an agent should act rationally he must not simultaneously pursue a goal g and a goal g' if g prevents the achievement of g'. Reasons for not pursuing some goals simultaneously are generally related to the fact that plans for reaching such goals may block each other [7].

Our proposal is based on argumentation-based reasoning, since it is a suitable approach for reasoning with inconsistent information [2]. The research questions that are addressed in this work are: (i) Can we identify when a kind of incompatibility arises by using arguments that represent the plans of the agent? if so, how would it be done? and (ii) faced with conflicting plans, how can the set of consistent goals be chosen?.

2. MOTIVATING EXAMPLE

The example scenario is based on the well-know "cleaner world" scenario. Two agents, BOB and TOM have the task of cleaning the dirt of an environment. There can be two kinds of dirt: liquid and solid. In the environment, there is also a workshop, where robots can go to be fixed or to recharge their batteries. The following situations may occur:

- ***Terminal incompatibility*:** Suppose that BOB detects dirt in a slot; then, the goal "cleaning slot" becomes pursuable. BOB also detects a minor technical defect in his antenna; hence, the goal "going to the workshop to be fixed" also becomes pursuable. BOB cannot commit to both goals at the same time because the plans adopted for each goal lead to an inconsistency, since he needs to be operative to clean the slot or non-operative to be fixed.

- ***Resource incompatibility*:** Suppose that BOB is in slot (2,4) and detects two dirty slots, slot (4,2) and slot (4,8). Therefore, goals "cleaning slot (4,2)" and "cleaning slot (4,8)" become pursuable. However, he only has battery for executing the plan of one of the goals. Consequently a conflict due to resource battery arises.

- ***Superfluity*:** Suppose that BOB is in slot (2,2) and he detects dirt in slot (4,4); hence, the goal "cleaning slot (4,4)" becomes pursuable. TOM also detects the same dirty slot and he also notes that it is liquid dirt; however, TOM's battery is low and he is not able to do the task, whereby he sends a message to BOB to mop slot (4,4); hence, the goal "mopping slot (4,4)" becomes pursuable. It is easy to notice that both goals have the same end, which is that slot (4,4) be cleaned, with the difference that "mopping slot (4,4)" is a more specific goal. This means that the agent will unnecessarily perform two plans that lead him to the same end. In a situation like this, we say that both plans are superfluous.

3. IDENTIFYING INCOMPATIBILITIES

We start by presenting the main mental states of the agent. The agent has finite sets of beliefs, goals (each goal has a preference value), resources, and actions, which are pairwise disjoint. The agent also has a finite of plans. We assume that all goals have at least one plan that allow the agent to achieve them.

*This abstract is a short version of the article "An Argumentation-based Approach for Identifying and Dealing with Incompatibilities among Procedural Goals", which is under review in the International Journal of Approximate Reasoning.

The process starts with a finite set of pursuable goals (possible incompatible). We represent the agent's plans by means of instrumental arguments and define the kinds of attacks that determine the incompatibilities.

The use of instrumental arguments for representing plans is not a novelty. Rahwan and Amgoud [5] define this kind of argument, which is structured like a tree where the nodes are rules whose components are desires and resources in the premise and a desire in the conclusion. Analogously, we use pursuable goals and resources in the premise and a pursuable goal in the conclusion. Additionally, we include beliefs and actions in the premise since this kind of argument represents an entire plan. We say that the set of rules in the nodes of the three are the support of the argument and the goal that is the conclusion of the rule in the root of the three is the claim of the argument. Figure 1 shows an example of instrumental arguments.

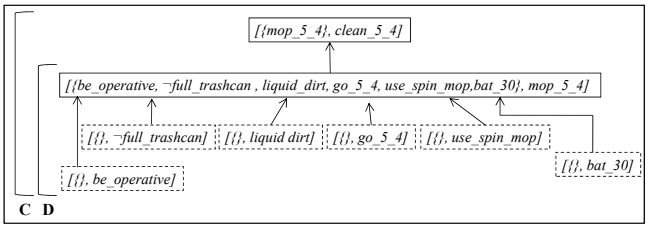

Figure 1: A tree-representation of two arguments. Goals *clean_5_4* **and** *mop_5_4* **are the claims of arguments** C **and** D**, respectively. Dotted-border squares represent the leaves of the tree.**

Regarding the attacks, we have identified one type of attack for each form of incompatibility:

- *Terminal incompatibility attack:* In this attack, the beliefs, the goals, and the actions of each rule of an argument are taken into account. We compare two arguments with different claims, which attack each other when their supports are logically inconsistent.
- *Resource incompatibility attack*: Two arguments are incompatible due to resources because the agent has no enough resources for performing the plans represented by both arguments; hence, we take into account the resources represented in each rule of the arguments.
- *Superfluity attack*: Superfluity emerges when two arguments have the same claim and different supports. Unlike other contexts, in practical reasoning the fact that two arguments support the same claim is considered unnecessary, or even worse, a waste of time or resources, because it means that the agent performs two plans when only one is necessary to achieve a given goal.

All of these attacks are symmetrical.

4. RESOLVING INCOMPATIBILITIES

We define a general argumentation framework (GAF) that includes all the arguments that can be generated from the sets of beliefs, goals, resources, and actions of the agent and all the attacks that arise between these arguments. We use the preference value of the goals in order to determine successful attacks, which breaks the symmetrical nature of attacks. Thus, an argument A successfully attacks another argument B when its preference value is greater than the preference value of B.

For resolving the incompatibilities and thus selecting the set of compatible goals, we use the so called argumentation semantics [2]. We propose to apply them in two different ways: (i) we apply a semantics on the instrumental arguments and the attacks that belong to the GAF, and (ii) we apply a semantics on a new argumentation framework – obtained from the GAF – that is made up of pursuable goals and the attacks between them. There is an attack between two pursuable goals g and g' when all the arguments whose claims are g attack – regardless the kind of attack – all the arguments whose claims are g'.

We propose to use only a conflict-free semantics. This means that we take sets of arguments, which does not attack. We made an analysis and concluded that the notions of defence and admissibility are not necessary in this problem. Over the conflict-free extensions, we filter the extensions that fulfill the two following criteria: (i) maximizing the number of compatible pursuable goals or (ii) maximizing the total preference value. Thus, from the conflict-free sets we take those that allow the agent to commit to the larger quantity of goals or those that allow the agent to commit to the goals that give him the highest utility.

Figure 2 shows a GAF where nodes represent arguments and edges attacks. After applying the conflict-free semantics and the two criteria aforementioned, the result is the extension $\{C, D, F\}$. Notice that arguments of goal *clean_5_4* (A and C) have a conflict with an argument of goal *be_fixed* (B); however, argument F does not conflict with A and C. Thus, both goals can be achieved. Goal *mop_5_4* is a subgoal of *clean_5_4*.

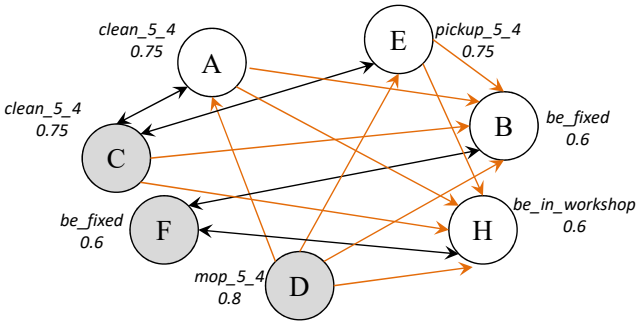

Figure 2: A GAF after considering the preference value. Orange arrows represent successful attacks. Next to each node there is information about the claim of the argument and the preference value of it. Gray coloured nodes represent the selected plans and therefore goals.

5. CONCLUSIONS

In this work, we have expressed a plan in terms of an instrumental argument and defined different kinds of attacks according to the form of incompatibility we wanted to identify. Both instrumental arguments and the attacks conform a general AF, on which it is a applied a conflict-free semantics in order to obtain the set of consistent goals. The main contributions of this work are: (i) the study and formalization of the resource incompatibility and the superfluity since these kinds of attacks has not been explored in the state of the art, and (ii) the two ways for selecting the goals the agent will commit to pursue.

REFERENCES

[1] C. Castelfranchi and F. Paglieri. The role of beliefs in goal dynamics: Prolegomena to a constructive theory of intentions. *Synthese*, 155(2):237–263, 2007.

[2] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial intelligence*, 77(2):321–357, 1995.

[3] I. Rahwan and L. Amgoud. An argumentation based approach for practical reasoning. In *Proceedings of the 5th International Joint Conference on Autonomous Agents and Multiagent Systems*, pages 347–354. ACM, 2006.

[4] M. B. van Riemsdijk, M. Dastani, and J.-J. C. Meyer. Goals in conflict: semantic foundations of goals in agent programming. *Journal on Autonomous Agents and Multi-Agent Systems*, 18(3):471–500, 2009.