Norm Adoption in the NoA Agent Architecture

Martin J. Kollingbaum
Department of Computing Science
University of Aberdeen
Aberdeen AB24 3UE, UK
Tel: +44 (0) 1224 272293
mkolling@csd.abdn.ac.uk

Timothy J. Norman
Department of Computing Science
University of Aberdeen
Aberdeen AB24 3UE, UK
Tel: +44 (0) 1224 272284
tnorman@csd.abdn.ac.uk

ABSTRACT
Agents based on reactive planning architectures use pre-specified plans as behaviour specifications. Normative agents are motivated by norms in their behaviour: obligations motivate them to act, prohibitions motivate them to refrain from certain actions, and permissions (or privileges) and capabilities specify the range of possible actions for such an agent. An important issue for normative agents is: under what circumstances is it appropriate for an agent to adopt a new set of obligations, prohibitions or permissions and what effect does it have on the agent’s normative state? In answering this question, a critical issue is whether or not this set of norms is consistent with the agent’s current normative state. Consistency of a set of norms is discussed in detail in respect to a reactive planning agent architecture – NoA, but, it is argued, provides useful insight to the problem of norm consistency in general and particularly where practical reasoning agents are concerned.

Categories and Subject Descriptors
I.2.11 [Artificial Intelligence]: Distributed Artificial Intelligence – coherence and coordination, languages and structure, multiagent systems.

General Terms
Agent Architecture

Keywords
Agent Architecture, Norms and Contracts.

1. INTRODUCTION
Normative agents are motivated by norms such as obligations, permissions and prohibitions in their behaviour. When such an agent wants to sign a contract with other agents, an important question to answer will be: is the set of norms declared in the contract consistent with the set of norms the agent currently holds? In general, this is not an easy question to answer. For example, an agent considering the adoption of an obligation to achieve \( p \) may already be prohibited from bringing about \( p \). A less obvious conflict would occur if all plans that could be chosen to achieve \( p \) have side-effects that conflict with one prohibition or another. In this case, the agent has to take into account not only the state \( p \) that, according to the norm, it is obliged to achieve, but also the various means at its disposal to achieve \( p \). The agent must look ahead, and if it is endowed with the ability to automatically construct plans to achieve its goals it must, effectively, search the space of possible plans to determine whether there exists one that fulfils the obligation without violating some other norm. Reactive planning agents provide a compromise between the flexibility of planning and the efficiency of purely reactive agents by requiring the designer to declare a fixed set of plan procedures, organised in a set of hierarchies, that are selected through a process of means-ends analysis. This representation of the capabilities of an agent serves to restrict the search for applicable plans. Thus, the complexity of the problem of searching for a plan is minimised, although at a concomitant loss of run-time flexibility. It is of interest to what extent reactive planning agents are able to perform automatic consistency checks of norms. For this investigation, the NoA normative agent architecture is used. NoA supports the implementation of practical reasoning agents, balancing flexibility and efficiency, and this one-step look-ahead is also adopted as a heuristic to determine the consistency of a norm.

2. NoA
The specific concern of the NoA agent architecture [1] is to provide means for the development of norm-motivated practical reasoning agents. The main elements that influence the behaviour of a NoA agent are (a) a set of beliefs, (b) a set of pre-specified plans and (c) a set of norms. These concepts are reflected in the NoA specification language. A NoA plan specifies when it is appropriate for execution (preconditions), what states of affairs it will achieve (effects) and the actual behaviour specification as the body of the plan. NoA plans have multiple effects, one of which will be the reason for its selection. Norm specifications carry activation and expiration conditions that determine when a norm becomes active and therefore relevant to an agent and when a norm ceases to be active. Obligations motivate a NoA agent to act through the selection of an appropriate plan. Prohibitions restrict the set of plans that can be selected. Permissions override prohibitions. NoA is based on the logic of responsibility for states and actions proposed by [3] and therefore captures the distinction between an agent taking responsibility for the achievement of a state of affairs and taking responsibility for the performance of an action (see [2]).

3. NORM ADOPTION
For an agent to adopt a new obligation, prohibition or permission, it should be consistent with extant norms. In other words, the introduction of a new obligation should not counteract existing obligations or motivate the achievement of a state of affairs that is forbidden. It is important to note that the effects of plans chosen for such achievements can produce such inconsistencies. In case of NoA, it therefore depends on the choice of plans if
inconsistencies occur. Not all plans would necessarily have effects that counteract existing norms. In the discussion below we refer to the effects of a plan $p$ as $\text{effects}(p)$ and to the negation of the set of effect of plan $p$ as $\text{neg\_effects}(p)$ (e.g. if $\text{effects}(p) = \{\text{on}(a, b), \text{not}\_\text{on}(a, c)\}$ then $\text{neg\_effects}(p) = \{\text{not}\_\text{on}(a, b), \text{on}(a, c)\}$).

Obligations, adopted under specific circumstances, can conflict with other norms. The term “strong inconsistency” determines these circumstances. A set of norms is strongly inconsistent, when there is necessarily a conflict. With the adoption of a new obligation, the state $s$ motivated to be achieved by this obligation is strongly inconsistent with $S_F$ (the set of states forbidden to be achieved) and with $S_O$ (the set of states obliged to be achieved) if and only if

$$\forall p \in \text{PLANS. If } s \in \text{effects}(p) \text{ then } \{S_F \cap \text{effects}(p) \neq \emptyset \text{ or } S_O \cap \text{neg\_effects}(p) \neq \emptyset\}$$

If a new obligation is added, then inconsistency occurs, if whatever plan is chosen to achieve the new state $s$, the plans produce effects that are forbidden or would counteract existing obligations. If a plan chosen for achieving state $s$ would produce a (side-) effect that is element of $S_F$ (set of forbidden states), then a strong inconsistency occurs. If an obligation from the set $S_O$ (set of obliged states) prescribes the achievement of a state and if an effect of the chosen plan is to achieve the negation of this state, then it is strongly inconsistent.

“Strong consistency” determines a situation where no conflict can occur. A set of norms is strongly consistent, when there is necessarily no conflict. With the adoption of a new obligation, the state $s$ motivated to be achieved by this obligation is strongly consistent with $S_F$ (the set of states forbidden to be achieved) and strongly consistent with $S_O$ (the set of states obliged to be achieved) if and only if

$$\forall p \in \text{PLANS. If } s \in \text{effects}(p) \text{ then } \{S_F \cap \text{effects}(p) = \emptyset \text{ and } S_O \cap \text{neg\_effects}(p) = \emptyset\}$$

If a new obligation is added, then consistency is guaranteed, if whatever plan is chosen to achieve the new state $s$, the plans do not produce effects that are forbidden and would counteract existing obligations. Strong consistency is guaranteed, if no plan chosen for achieving state $s$ would produce a (side-) effect that is element of $S_F$ (set of forbidden states) and if no plan chosen for achieving state $s$ has an effect to not achieve a state that is element of $S_O$ (set of obliged states).

“Weak consistency” determines a situation where the addition of an obligation possibly leads to an inconsistency. A set of norms is weakly consistent, when there is possibly a conflict. With the adoption of a new obligation, the state $s$ motivated to be achieved by this obligation is weakly consistent with $S_F$ (the set of states forbidden to be achieved) and weakly consistent with $S_O$ (the set of states obliged to be achieved) if and only if

$$\exists p \in \text{PLANS s.t. } s \in \text{effects}(p) \text{ and } \{S_F \cap \text{effects}(p) = \emptyset \text{ and } S_O \cap \text{neg\_effects}(p) = \emptyset\}$$

As long as there is at least one plan that can be chosen to achieve state $s$, which does not produce any of the forbidden states as an effect and does not counteract any obligations, weak consistency occurs.

Prohibitions are not motivators for activities, they introduce a restriction on the behaviour of an agent. When a prohibition is adopted, inconsistencies with existing obligations can occur. The adoption of a prohibition may have one of the following outcomes: (a) the set of norms remains in its current state, which can be strong / weak consistency or strong inconsistency, or (b) the set of norms is transferred or “downgraded” into a weaker state of consistency, which can be weak consistency or strong inconsistency.

Permissions override prohibitions. Therefore, they “upgrade” the consistency state of a set of norms. The adoption of a permission may have one of the following outcomes: (a) the set of norms remains in its current state, which can be strong / weak consistency or strong inconsistency, or (b) the set of norms is transferred or “upgraded” into a stronger state of consistency, which can be weak / strong consistency.

4. CONCLUSION
In this paper, we have focused on issues of consistency of sets of norms adopted by a practical reasoning agent. Of particular interest is the “level of consistency” that occurs if an agent adopts a new norm. In this paper, three levels of consistency are described: “strong consistency”, “weak consistency” and “strong inconsistency”. An obligation may be strongly inconsistent, weakly consistent or consistent with a set of norms (obligations, permissions and prohibitions). Similarly the consistency of a prohibition will be at one of those levels. Finally, permissions have the potential to increase the consistency of the set of norms governing an agent’s behaviour. Using this classification of levels of consistency, an agent may make reasonable decisions on whether or not to adopt a new norm. With the ability to represent and reason about and act according to norms, the NoA agent architecture provides a solution for automated contract management and a technology for the development and deployment of norm-governed agents.

REFERENCES