Combining Goal Generation and Planning in an Argumentation Framework

Joris Hulstijn[a] Leendert van der Torre[b]

[a] Universiteit Utrecht
[b] CWI Amsterdam

Abstract

In this paper we use formal argumentation theory to study the deliberation cycle of the BOID cognitive agent architecture. We combine goal generation with a planning approach recently proposed by Leila Amgoud. In her formal argumentation framework plans for fulfilling a desire are represented just like arguments that support a conclusion; the framework can thus deal with conflicting desires. In this paper we show how the argumentation approach to planning can be combined with goal generation in the BOID architecture. We also show that combining goal generation and planning leads to interesting new research questions about their relationship.

1 Introduction

An important aspect of intelligent agents is the goal deliberation process, in which an agent decides which goals to pursue. Goal deliberation can be implemented in many ways, though agent architectures often assume separate components for goal generation, goal selection and planning [9, p 76]. Terminology differs; we say that a goal generation process produces potential goals on the basis of an agent's beliefs, desires (internal motivation), and possibly obligations (external motivation). Goal selection is the subsequent process of deciding which consistent subset of goals will be pursued. These selected goals become intentions. Planning is the process of finding a sequence of actions to achieve the intentions.

In this paper we study goal deliberation in the BOID cognitive agent architecture [2, 3, 5]. Goal generation in the BOID architecture has been studied before, but planning is the subject of a joint project with the IRIT laboratory in Toulouse, which is developing an argumentation theoretic account of planning [1].

The BOID architecture uses sets of prioritized default rules to specify the contents of four separate components: Belief, Obligation, Intention and Desire. As in most rule-based systems, heads of rules are compared to the facts, producing a set of applicable rules. Based on a priority order, a rule is selected and applied. Earlier we showed how restrictions on priority orders correspond to agent types [3]. For example, a selfish agent ranks desire rules over obligation rules.

In goal deliberation we have to deal with two kinds of reasoning: forward reasoning from current beliefs to desirable states (deduction), and backward "means-ends" reasoning from desirable states to required actions (abduction). This combination is often problematic. We have observed two problems with goal deliberation in the BOID architecture.

1. The goal generation and goal selection components partly depend on planning, because potential or selected goals are required to be feasible [4, 7]. A goal is feasible when some plan exists that is likely to achieve it. Rao and Georgeff call an agent that only generates feasible goals *strongly realistic*. It is waste of resources to consider infeasible goals. The feasibility restriction requires that planning can provide feedback to goal generation and selection.
2. The application of a priority order is difficult. The easiest solution is to define a local priority order over rules [8]. However, single rules often have undesired consequences. Application of rules should therefore be compared by their outcomes, using a utility value for example. In a possible worlds framework utilities can be expressed, but in a practical agent system this is difficult to achieve. In previous work we have chosen to use extensions, maximally consistent sets of rules, as the main interface between components [2, 3, 5]. The use of extensions has some drawbacks, both practical and conceptual. The extensions themselves become large, and their consistency becomes difficult to check and maintain. The number of extensions multiplies quickly, while the overlap between extensions is considerable. Clearly, we need a clever representation.

To address these problems, we apply techniques from argumentation theory [6]. We combine goal generation with Amgoud's version of Dung's argumentation framework to reason about conflicting desires [1]. The central analogy is the following. A desire or potential goal that has possible 'trees of realization' or plans to achieve it, can be modeled just like an argument which consists of a conclusion with the supporting argumentations. Trees of realization form a declarative representation at the intermediate level between rules and outcomes. The attack relation defined over arguments can serve as a criterium to select goals. Because infeasible goals have no tree of realization, they are excluded from deliberation.

In this paper we use formal argumentation theory to study the deliberation cycle of the BOID architecture. Due to space limitations, we do not discuss the advantages and disadvantages of argumentation theory with respect to theories developed in, for example, conditional logic, logic programming or databases. We just observe that argumentation and dialogue theories have been used to analyze logic, and that defeasible argumentation has been used to study non-monotonic reasoning and reasoning about uncertainty. Theories for conditionals and rules have attracted attention lately with dedicated workshops, and a further unification may be expected during the coming years.

The paper is structured as follows. In section 2 we present goal generation, in section 3 we combine it with Amgoud's argumentation framework, and in section 4 we consider new research questions.

2 BOID goal generation

To keep discussion to a minimum, we only consider the generation of goals from desires, not from obligations and intentions. In the language, we distinguish among proposition variables that we call decision variables and other variables. Decision variables represent atomic actions that need no further plan to be achieved. This distinction is implicit in Amgoud's framework.

Definition 1 (Logic) *Let A be a set of decision variables. Let N be a set of non-decision variables. Let L be a propositional language built from $A \cup N$. Let a literal l of L be a variable or its negation. Let a rule be an ordered nonempty finite list of literals written like $l_1 \wedge l_2 \wedge \ldots \wedge l_{n-1} \rightarrow l_n$. We call $l_1 \wedge l_2 \wedge \ldots \wedge l_{n-1}$ the body of the rule, and l_n the head. If $n = 1$ the body is empty and we write l_n.*

We extend Amgoud's notion of unconditional desires to a set of desire rules D, and we interpret a plan library and background knowledge as sets of rules P and Σ respectively. A desire rule $l_1 \wedge \ldots \wedge l_{n-1} \rightarrow l_n$ in D represents that l_n is desired in the context $l_1 \wedge \ldots \wedge l_{n-1}$, a planning rule in P represents that l_n is achieved if $l_1 \wedge \ldots \wedge l_{n-1}$ is achieved, and a background rule in Σ represents that l_n is true when $l_1 \wedge \ldots \wedge l_{n-1}$ is true. Since a decision variable needs no planning rules, we require that the head of a planning rule is not a decision variable.

Definition 2 (Desire-plan description) *Let A, N and L be as defined in definition 1. A* desire-plan description *is a tuple $\langle D, P, \Sigma \rangle$ with D, P and Σ sets of rules from L, such that the heads of rules in P are built from a variable in N.*

Here are some examples of desire-plan descriptions. The first example extends example 1 of [1]. Note that the examples in this paper are meant to illustrate the definitions. We realize that further validation of the approach is necessary.

Example 1 (Travel) *Assume variables $A = \{ag, fr, hop, dr, sol, w, cp, gp\}$ and $N = \{wca, jca, hjor, dlc, fp, pa, t, vac\}$, with the interpretation:*

wca	:	*there is a war in Central Africa*	*ag*	:	*to go to the agency*
jca	:	*to journey to Central Africa*	*fr*	:	*friend brings the tickets*
hjor	:	*have a job on return*	*hop*	:	*to go to a hospital*
dlc	:	*deadline for submission is close*	*dr*	:	*to go to a doctor*
fp	:	*to finish a paper before going*	*sol*	:	*to solicit for work*
pa	:	*the paper is accepted*	*w*	:	*to work*
t	:	*to get the tickets*	*cp*	:	*call the program chair*
vac	:	*to be vaccinated*	*gp*	:	*write a good paper*

Consider the following desire-plan description:

$$D = \{\neg wca \rightarrow jca,\ jca \rightarrow hjor,\ dlc \rightarrow fp,\ fp \rightarrow pa\}$$

$$P = \{t \wedge vac \rightarrow jca,\ ag \rightarrow t,\ fr \rightarrow t,\ hop \rightarrow vac,\ dr \rightarrow vac,\ sol \rightarrow hjor,\ w \rightarrow fp,\ cp \rightarrow pa,\ gp \rightarrow pa\}$$

$$\Sigma = \{\neg wca,\ dlc,\ w \rightarrow \neg ag,\ w \rightarrow \neg dr,\ w \rightarrow \neg hop\}$$

In some contexts, the agent has desires to travel, have a job, to finish a paper and get this paper accepted. There are several ways to achieve these desires. For example, to get a paper accepted this agent can either write a good paper or call the program chair.

The second example is concerned with food and wine.

Example 2 (Dinner) *Let $A = \{e, r, d, t, w\}$, $N = \emptyset$ and represent dinner options with e for entrecote, r for red whine, d for daurade, t for trout and w for white wine. Let dinner preferences be $D = \{e,\ e \rightarrow r,\ d,\ t,\ d \wedge t \rightarrow w\}$, $P = \emptyset$ and $\Sigma = \emptyset$. The agent prefers to have red wine with red meat, and white wine with fish.*

A goal set is a set of related desires. First we define what it means to apply desire rules to generate a goal set, which is analogous to applying inference rules in classical logic. Then we define a goal set to be a set of heads of desire rules that cannot be further split up into separate goal sets. This represents the fact that desires in a goal set are related. The definition may seem circular, but note that the size of the goal sets shrinks and that they are finite. The notion is thus well-defined. Finally we define maximal goal sets with respect to set inclusion.

Definition 3 (Goal set) *The closure of a set of rules R on a set of literals S, written as $C(R, S)$, is defined by $C(R, S) = \bigcup_{i=0}^{i=\infty} S^i$ with $S^0 = S$ and $S^{i+1} = S^i \cup \{l \mid l_1 \wedge \ldots \wedge l_n \rightarrow l \in R \text{ and } \{l_1, ..., l_n\} \subseteq S^i\}$. We write $H(R)$ for the set of heads of the rules in R. Given a desire-plan description $\langle D, P, \Sigma \rangle$, a finite set of literals GS is a* goal set *iff there exists a subset D' of D such that:*

1. *$GS = C(D' \cup \Sigma, \emptyset) \cap H(D')$;*
2. *$C(D' \cup \Sigma, \emptyset)$ is consistent, i.e., does not contain two literals l and $\neg l$;*
3. *There is no set of goal sets $\{GS_1, \ldots, GS_n\}$ with each $GS_i \neq GS$ and $GS = GS_1 \cup \ldots \cup GS_n$.*

A maximal goal set *is a goal set which is maximal with respect to set inclusion.*

The following two examples illustrate that a goal set is a set of *related* desires.

Example 3 (Travel, continued) *The goal sets are $\{jca\}$, $\{jca, hjor\}$, $\{fp\}$ and $\{fp, pa\}$. The set of desires with $\{jca,\ hjor,\ fp,\ pa\}$ is not a goal set, because it can be split in $\{jca, hjor\}$ and $\{fp, pa\}$. The latter two are the maximal goal sets. Here, jca and hjor are related, because the desire to have a job on return from travel (hjor) is conditional on making a journey to Central Africa (jca). Likewise, (fp) and (pa) are related, because the desire to have a paper accepted (pa) is conditional on finishing a paper before going to Central Africa (fp).*

Example 4 (Dinner, continued) *We have the following goal sets: $\{e\}$, $\{e, r\}$, $\{d\}$, $\{t\}$, $\{d, t, w\}$. The set of desires $\{e, r, d, t, w\}$ is not a goal set, because it can be split in $\{e, r\}$ and $\{d, t, w\}$. This expresses for example that the desires for e and r are related, but that desires for e and d are not related. The sets $\{e, r\}$ and $\{d, t, w\}$ happen to be the maximal goal sets. Here, e and r are related, because the desire for red wine (r) is conditional on having entrecote (e). Likewise, the desire for white wine (w) is conditional on having daurade (d) and trout (t) for dinner.*

3 Amgoud's framework for planning

In this paper we combine Amgoud's definitions of an argumentation framework with our goal sets. Her motivation to study plan generation in an argumentation framework is to deal with conflicting desires. She also shows how plan argumentation differs from belief argumentation. Actions and sub-actions are defined as tuples $\langle h, H\rangle$ analogous to claims for h with support H [1, 6].

Definition 4 (Action) *An* action *for a desire-plan description* $\langle D, P, \Sigma\rangle$ *is:*
- $\langle h, \emptyset\rangle$ *for any decision variable* h*, called an atomic action; or*
- $\langle h, \{l_1, \ldots, l_n\}\rangle$ *for any rule* $l_1 \wedge \ldots \wedge l_n \rightarrow h \in P \cup \Sigma$.

Two actions $\langle h_1, H_1\rangle$ *and* $\langle h_2, H_2\rangle$ *conflict iff* $C(\Sigma, \{h_1, h_2\} \cup H_1 \cup H_2) \vdash \bot$.

From these actions Amgoud constructs so-called realization trees, which serve as a way to represent and reason about plans. Each node of the tree is an action and each child of an action is one its subactions. We extend Amgoud's definition of a realization tree for goal sets, in which the root contains a set of desires. Realization trees are plans to achieve a particular goal set, and may thus also be called goal achievement trees.

Definition 5 *A* tree of realization g *for a goal set* GS*, written as* $g(GS)$*, is a finite tree consisting of actions such that*
- $\langle \top, GS\rangle$ *is the root of the tree;*
- *A node* $\langle h, \{l_1, \ldots, l_n\}\rangle$ *has exactly* n *children* $\langle l_1, H_1\rangle, \ldots, \langle l_n, H_n\rangle$*;*
- *The leaves of the tree are atomic actions.*

Consider the holiday plans of example 1. Figure 1 visualizes two realization trees. One for the goal set $\{jca, hjor\}$, which is based on desire rules $\neg wca \rightarrow jca$ and $jca \rightarrow hjor$, and one for $\{pa, fp\}$, based on desire rules $dlc \rightarrow fp$ and $fp \rightarrow pa$. Note that the goal generation steps are not visualized.

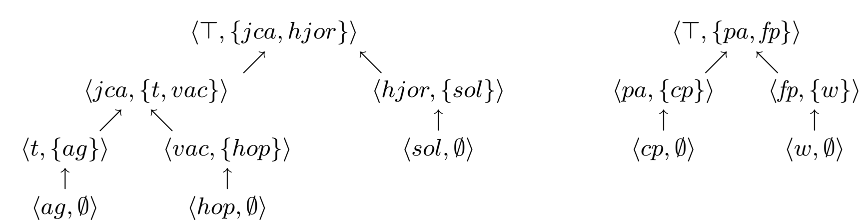

Figure 1: Two trees of realization

Just like arguments, trees of realization may conflict. Therefore it makes sense to use concepts from argumentation theory. In general, an argumentation framework is defined as a set of arguments with a binary relation that represents which arguments attack which other arguments [6]. Here, trees of realization play the role of arguments; the attack relation is derived from conflicts between actions. For more details on this argumentation framework, we refer to Amgoud's paper [1].

Definition 6 *A* system handling conflicting desires *(*SHD*) is a tuple* $\langle G, Attack \rangle$ *such that* G *is a set of realization trees and* $Attack$ *is a binary relation over* G*.*

Based on the notion of conflict between actions (definition 4), we can specify a particular attack relation. The argumentation notions of *defence*, *preferred extension* and *basic extension* are defined accordingly. Note that Amgoud's definition of defence differs from the usual one.

Definition 7 *Let* $\langle G, Attack \rangle$ *be an* SHD *such that* G *contains all realization trees that can be derived from goal sets of a given desire-plan description* $\langle D, P, \Sigma \rangle$*. Let* $S \subseteq G$ *and* $g, g_1, g_2 \in G$ *be (sets of) realization trees.*

- $\langle g_1, g_2 \rangle \in Attack$*, i.e.,* g_1 *attacks* g_2*, iff there exist actions* a_1 *and* a_2 *in the nodes of* g_1 *and* g_2 *respectively, such that* a_1 *and* a_2 *conflict.*
- S *is attack free iff there are no* $g_1, g_2 \in S$ *such that* g_1 *attacks* g_2*.*
- S *defends* g *iff for all* $g_1 \in G$ *such that* g_1 *attacks* g*, there is an alternative* $g_2 \in S$ *with* $root(g_1) = \langle \top, GS \rangle$ *and* $root(g_2) = \langle \top, GS \rangle$*, for some* GS*.*
- S *is a preferred extension iff* S *is maximal w.r.t. set inclusion among the subsets of* G *that are attack free and that defend all their elements.*
- S *is a basic extension iff it is a least fixpoint of the function* $F(S) = \{g | g \text{ is defended by } S\}$*.*

These argumentation notions are illustrated by the travel example.[1]

Example 5 (Travel, continued) *There are 7 realization trees. Two for goal set* $\{jca\}$*, one with hop and one with dr, and therefore also two for* $\{jca, hjor\}$*. There is one tree for* $\{fp\}$ *and again two for* $\{fp, pa\}$*, one with cp and one with gp. The two trees in figure 1 attack each other, amongst other reasons because actions* $\langle ag, \emptyset \rangle$ *and* $\langle w, \emptyset \rangle$ *conflict. In fact any tree for* $\{jca, hjor\}$ *or* $\{jca\}$ *attacks any tree for* $\{fp, pa\}$ *or* $\{fp\}$*. These two clusters of trees are internally attack free. There is no non-trivial defend relation in this example. The preferred extensions are the set of trees for the cluster* $\{jca, hjor\}$ *or* $\{jca\}$*, and the set of trees for the cluster* $\{fp, pa\}$ *or* $\{fp\}$*. The basic extensions are the same.*

4 Intertwining goal generation and planning

Thus far, the generation of goal sets and of realization trees (planning) have been largely kept separate. However, we can extend realization trees (plans) with a representation of goal generation steps. To distinguish goal generation steps from planning steps, we let the arrow point in the opposite direction. Consequently the representation is no longer a tree and will from now on be called a *directed acyclic graph* or DAG. The arrows indicate the flow of reasoning in the following sense. A downward arrow for goal generation represents a deduction step. In example 1, from the desire jca and $jca \rightarrow hjor$ we may conclude $\{jca, hjor\}$. An upward arrow indicates an abduction step. For example, from jca we can abduce $\{t, vac\}$ with the rule $t \wedge vac \rightarrow jca$. The following example illustrates that in some cases, goal generation and planning are even more intertwined.

[1] The definitions and examples are implemented in Prolog, see `http://boid.info/boidarg/`.

Example 6 *Let* $A = \{a, b\}$ *and* $N = \{p, q\}$. *Consider the desire-plan description* $D = \{\top \rightarrow p, a \rightarrow q\}$, $P = \{a \rightarrow p, b \rightarrow q\}$, $\Sigma = \emptyset$. *An example of a desire for* q *triggered by* a, *is the desire to drink during extended physical exercise, such as a rowing race. If we extend the formalism to obligations (external motivation), another example would be the obligation to pay on the metro. The obligation is triggered by the traveling action; it is not a precondition in the usual sense. In this framework, preconditions or required resources are best represented by a planning rule. The only goal set is* $\{p\}$, *and the only tree of realization for* p *contains the actions* $\langle p, \{a\}\rangle$ *and* $\langle a, \emptyset\rangle$. *However, intuitively the action* $\langle a, \emptyset\rangle$ *seems to trigger a further desire for* q. *When we extend a tree of realization with goal set generation, we get the* DAG *on the left of figure 2.*

Example 7 *Let* A *and* N *be as in example 6. Consider* $D = \{\top \rightarrow p, a \rightarrow q\}$, $P = \{a \wedge b \rightarrow p, a \wedge b \rightarrow q\}$, $\Sigma = \emptyset$. *This* DAG *is shown at the right of figure 2.*

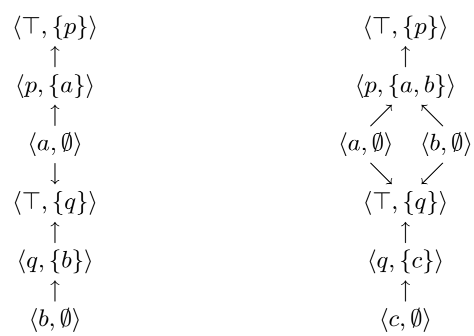

Figure 2: Dags of realization with alternation for examples 6 and 7

5 Concluding remarks

In this paper we have combined BOID goal generation with planning in an argumentation framework. As a result, argumentation notions like attack, defend, maximally preferred extensions and basic extensions, can be used in a model of goal deliberation. Reconsider the two problems discussed in the introduction.

1. The preferred extension corresponds to the maximal set of mutually consistent feasible goals: for each goal, there is a plan and for all possible conflicts between plans, there is an alternative. Similarly, the basic extension can be interpreted as a stable set (self-defending) of feasible goals. Therefore such goal sets are good candidates to become intentions.
2. We suggest to use argumentation extensions, which are sets of realization trees. Realization trees provide an intermediate structure to allow for consistency checks and comparison. Goals without a defended realization tree can be filtered out.

The use of realization trees thus promises to provide a representation layer at a level between individual rules and complete outcomes. Because an infeasible goal does not have a defended tree of realization, the feasibility restriction is easily applied.

Dags of realization promise a tighter connection between goal generation and planning. In particular, they are able to express alternating goal generation and planning steps, illustrated by example 6. This possibility illustrates that more research on the nature of the relationship between goal generation and planning is required. In future work we will consider ways of extending the argumentation notions to dags of realization. Dags of realization form a representation that addresses feedback between planning and goal generation. However, such a representation format only provides a first step; more cognitive and applied research into the nature and desirability of feedback between goal generation and planning remains necessary.

Acknowledgements

Thanks to Leila Amgoud for discussions on the topics raised in this paper when the second author visited the IRIT laboratory in March 2003. Thanks also to Mehdi Dastani for valuable input.

References

[1] L. Amgoud. A formal framework for handling conflicting desires. In *Proceedings of Seventh European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU 2003)*, LNCS 2711, 552-563, Springer Verlag, 2003.

[2] J. Broersen, M. Dastani, Z. Huang, J. Hulstijn, and L. van der Torre. The BOID architecture. In *Proceedings of the Fifth International Conference on Autonomous Agents (AA'2001)*, 9-16, ACM Press, 2001.

[3] J. Broersen, M. Dastani, J. Hulstijn, and L. Van der Torre. Goal generation in the BOID architecture. *Cognitive Science Quarterly*, 2(3-4):431–450, 2002.

[4] P.R. Cohen and H.J. Levesque. Intention is choice with commitment. *Artificial Intelligence*, 42:213–261, 1990.

[5] M. Dastani and L. van der Torre. What is a normative goal? Towards goal-based normative agent architectures. In *Proceedings of International Workshop on Regulated Agent-Based Social Systems: Theories and Applications (RASTA'02)*. Springer, to appear.

[6] P. M. Dung. On the acceptability of arguments and its fundamental role in non-monotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77:321–357, 1995.

[7] A. S. Rao and M. P. Georgeff. Decision procedures for BDI logics. *Journal of Logic and Computation*, 8(3):293–343, June 1998.

[8] R. Reiter. A logic for default reasoning. *Artificial Intelligence*, 13:81–132, 1980.

[9] M. Wooldridge. *An Introduction to MultiAgent Systems*. Wiley, 2002.