Defeasible Goals

Leendert van der Torre

Max Planck Institute for Computer Science
Im Stadtwald, D-66123 Saarbrücken, Germany

torre@mpi-sb.mpg.de

Abstract. In this paper we study defeasible goals in Gabbay's labelled deductive systems. We prove the completeness of a simple and elegant proof theory for the labelled logic of defeasible goals by proving two phasing theorems.

1 Introduction

In the usual approaches to planning in AI, a planning agent is provided with a description of some state of affairs, a *goal state*, and charged with the task of discovering (or performing) some sequence of actions to achieve that goal. Recently several *logics* for goals (and the closely related desires) have been proposed [4, 3, 12, 2, 14, 13, 8, 6, 15, 16]. We are interested in a logical system that tells us which conditional goals can be derived from a set of conditional goals (the set of premises) called the goal base. In this simple setting – in which we do not consider logical connectives between the conditional goals – reasoning about goals is non-trivial for the following three reasons.

Goals can conflict. Some students desire to do research and desire to be rich, although they realize that they cannot achieve both goals simultaneously.

Goals are context dependent. If the robot's goal to get me some coffee cannot be achieved, for example because the coffee machine is broken, then its alternative goal may be to get me some tea. In preference-based approaches this goal is a preference of tea without coffee over neither tea nor coffee.

Goals can be overridden by other goals. Bacchus and Grove's [1] surgery and marriage examples below illustrate that the addition of a goal to the goal base may result in new goals coming into force and, more importantly, old goals losing their force.

Example 1: surgery. A person may prefer not having surgery over having surgery, but this preference might be reversed in the circumstances where surgery improves one's long term health. We can defeasibly infer that the person prefers no surgery only as long as it is not known that surgery improves his or her long term health.

Example 2: marriage. A girl named Sue prefers to be married to John over not, she prefers to be married to Fred over not, and at the same time she reasonably prefers to be married to neither over being married to both. We can defeasibly infer that she likes to be married to both from the first two preferences, but not from all three together.

The conditional goal 'α is preferred (or desired) if β' is represented by the formula $G(\alpha|\beta)$. Desirable properties of the logic of conditional goals are strengthening of the antecedent to deal with irrelevant circumstances, weakening of the consequent to support reasoning about sub-goals, the conjunction rule for the consequent and cumulative transitivity (or cut) to combine goals, and the disjunction rule for the antecedent to support reasoning by cases. Moreover, in contrast to most conditional and counterfactual logics the identity $G(\alpha|\alpha)$ is an undesirable property, because if α is the case then it does not follow that there is a goal for α. The examples above illustrate that intuitively most proof rules of the logic of *defeasible* goals – i.e. goals which can be overridden by other goals – only hold in a restricted sense. This is in line with negative results obtained in the logic of preference, where counterexamples have been given to nearly all proposed rules [11]. In semantic terms the problem of the logic of preference is that there is no straight forward way to lift a preference relation between worlds, for example represented by a utility function, to a preference relation between sets of worlds or propositions. Pearl [12] also observed counterexamples to most proof rules when utilities and probabilities are combined in a so-called decision-theoretic logic of preference. We study restricted applicability with a global consistency constraint on proofs and with a local consistency constraint with phasing, as explained below. The main result of this paper is that the two approaches are equivalent.

First, the restricted applicability of proof rules is enforced by a *consistency constraint* on proofs [10]. For example, if we ignore complications raised by reasoning by cases and overriding, then a proof is only valid if the set of the fulfillments of the goals at the nodes of its proof tree is consistent. Here a proof is a proof tree with – due to the restricted language – only conditional goals at its nodes, and the fulfillment of the goal $G(\alpha|\beta)$ is the formula $\alpha \wedge \beta$. This is a global constraint on proofs in the sense that it cannot be checked locally *at each derivation step* whether the constraint is violated or not. The consistency constraint blocks derivations from conflicting goals (consistency check on consequents), as well as derivations from goals with conflicting contexts (consistency check on antecedents). Moreover, in the logic of defeasible goals introduced in this paper the consistency check not only covers the premises a goal is derived from, but also other goals of the goal base it depends on. To formalize this dependence on the goal base we define an inference relation relative to the goal base, as is explained in detail later. With this extension we do not only cover conflicts and context dependence, but also the overriding of the surgery and marriage examples.

Second, the restricted applicability of proof rules is enforced by restricting the *order of application* of the proof rules, together with a local consistency constraint at each derivation step. If we ignore the local consistency constraint and replacements by logical equivalents, then the proof theory is as follows. The goal $G(\alpha|\beta_1 \wedge \beta_2)$ is *an atomic goal* of goal base B if there is a goal $G(\alpha|\beta_1) \in B$, the goal $G(\alpha_1 \wedge \ldots \wedge \alpha_n|\beta)$ is *a compound goal* of B if there is a set of atomic goals $\{G(\alpha_1|\beta), G(\alpha_2|\beta \wedge \alpha_1), \ldots, G(\alpha_n|\beta \wedge \alpha_1 \wedge \ldots \wedge \alpha_{n-1})\}$ of B and the goal $G(\alpha|\beta_1 \vee \ldots \vee \beta_n)$ is *a combined goal* of B if there is a set of compound goals

$\{G(\alpha_1|\beta_1), \ldots, (G(\alpha_n|\beta_n)\}$ of B such that $\alpha_i \vdash \alpha$ for $i = 1 \ldots n$. Atomic goals are derived from the goal base by strengthening of the antecedent, compound goals are derived from atomic goals by cumulative transitivity and the conjunction rule, and combined goals are derived from compound goals by weakening of the consequent and the disjunction rule. Hence, this is the only admissible order of proof rule application.

The theorems of this paper imply that a goal can be derived from the goal base B with the global consistency constraint if and only if it is a combined goal of B. To prove these surprising theorems the global consistency constraint as well as the ordering of proof rules is implemented in versions of Gabbay's labelled deductive systems [5]. In particular, we use and adapt the labelled logics introduced by Van der Torre [15, 16] and Makinson [9] for defeasible goals. The labelled logic of conditional goals is based on a language of formulas of the type $G(\alpha \mid \beta)_L$, to be read as 'α is preferred (or desired) if β in the proof context L.' The label L collects the goals used in the proof to check the consistency of the fulfillments, and it collects the proof rules used in the proof to check the admissible orders. The formalization in a labelled deductive system has two advantages. First, as shown in [16], we can use rewriting techniques for two subsequent derivation steps, and generalize them inductively to rewrite arbitrary proof trees. Second, as shown in this paper, for the logic of defeasible goals we can reuse results of the logic of non-defeasible goals by proving a formal relation between the two logics.

This paper is organized as follows. In Section 2 we introduce the logic, and in Section 3 we illustrate the logic by several examples. In Section 4 we prove the relation between this new logic PLLDG and its predecessor PLLG, which is the basis of the completeness proof of the phased proof theory in Section 5.

2 Phased labelled logics of defeasible goals (PLLDG)

The first labelled logic of the type considered in this paper was proposed in [17] for obligations in a normative context. In this rudimentary logic the label of a premise only contains the consequent, and the consistency check considers the label together with the antecedent of the potentially derived goal. Two extensions of this logic have been proposed to cover the parallel tracks created through reasoning by cases, such that $G(a|\top)$ can be derived from the conflicting goals $G(a \wedge b|c)$ and $G(a \wedge \neg b|\neg c)$. Here $\top$ stands for any tautology like $p \vee \neg p$. First, in [15] the so-called violation restriction is based on the material conditional $\beta \rightarrow \alpha$ of premises $G(\alpha \mid \beta)$ (the negation of the material conditional $\beta \rightarrow \alpha$ is the violation $\beta \wedge \neg\alpha$), and it is suggested that this restriction is best suited for normative contexts.[1] Second, in [9] the label contains a separate set of goals

[1] In [15] the violation restriction checks whether fulfilling a derived goal does not imply the violation of one of the goals it is derived from. However, to block some obvious counterintuitive derivations in Example 2 of that paper this should be strengthened to the restriction that fulfilling a derived goal does not imply the disjunction of the violations of the goals it is derived from.

for each case. The OR rule is thought of as creating different cases or parallel dependency tracks. Given a node in a derivation tree, we travel upwards the derivation tree generated by the node, splitting in two cases at every application of the OR rule, such that one case contains one of the two premise nodes and the other case contains the other premise node. As far as reasoning by cases is concerned the consistency check can either be based on material implications $\beta \rightarrow \alpha$, consequents α only, or fulfillments $\alpha \wedge \beta$. In this paper we study the latter most cautious option.

Definition 1 (Cases). *Consider a tree-like proof of a goal (the root) from a set of premises (the leaves). A case or dependency tree of a proof is represented by a set of nodes from the tree. The set of cases of the proof can be computed with the following procedure.*

1. *Initially, the set C contains one element, the set with the root.*
2. *As long as there is a non-leaf node in some set of C without one of its parents, then either:*
   - *if the node is not derived by* OR*, then add its parents to the set;*
   - *otherwise create two copies of the set, one for each of its parents; add this parent.*
3. *Each set of nodes in C represents a case of the proof.*

For example, consider the proof tree below of the derivation of $G(a|\top)$ from $G(a \wedge b|c)$ and $G(a \wedge \neg b|\neg c)$. There are two cases, $\{G(a \wedge b|c), G(a|c), G(a|\top)\}$ and $\{G(a \wedge \neg b|\neg c), G(a|\neg c), G(a|\top)\}$, and the sets of fulfillments of both cases are consistent.

$$\dfrac{\dfrac{G(a \wedge b|c)}{G(a|c)}\,\text{WC} \qquad \dfrac{G(a \wedge \neg b|\neg c)}{G(a|\neg c)}\,\text{WC}}{G(a|\top)}\,\text{OR}$$

For non-defeasible goals [16] a consistency check ensures that a goal may only be derived if the set of fulfillments of goals of each of its cases is consistent. It must always be possible to fulfill a derived goal together with each of the goals it is derived from, though – to support reasoning by cases – not necessarily all cases at the same time. In this paper we study labelled logics of *defeasible* goals, in which goals can be overridden by other dependent goals to resolve conflicts. The consistency check therefore also ensures that a derived goal does not violate a dependent goal. Here 'dependence' is formalized by an explicit relation (I) and 'violation' is defined as follows.

Definition 2 (Goal violation). *A goal $G(\alpha|\beta)$ derived from $G(\alpha'|\beta')$ violates the goal $G(\alpha_1|\beta_1)$ iff*

- $\alpha' \wedge \beta' \wedge (\beta_1 \rightarrow \alpha_1)$ *is classically consistent and*
- $\alpha \wedge \beta \wedge (\beta_1 \rightarrow \alpha_1)$ *is classically inconsistent.*

For example, the derivation of $G(p|q)$ from $G(p|\top)$ violates the goal $G(\neg q|\top)$, because $\{p, \neg q\}$ is consistent whereas $\{p \wedge q, \neg q\}$ is not. Moreover, the derivation of $G(j \wedge f|\top)$ from $G(j|\top)$ and $G(f|\top)$ violates the goal $G(\neg(j \wedge f)|\top)$.

To facilitate the definitions, we do not collect all goals of a proof tree in the label, but only the premises together with so-called strengthenings, where the strengthening β_2 derives $G(\alpha|\beta_1 \wedge \beta_2)$ from $G(\alpha|\beta_1)$ by applying the proof rule strengthening of the antecedent. It can easily be shown by an inductive argument that the set of fulfillments of premises together with strengthenings is equivalent to the set of fulfillments of all goals. Moreover, instead of collecting the proof rules explicitly we associate an integer with each proof rule called its phase, and only represent this integer in the label [16]. Each formula occurring as a premise has a label that consists of itself and phase 0. The phase of a goal is determined by the proof rule used to derive the goal, and the labels are used to check that the phase of reasoning is non-decreasing. Note that the inference relation $\vdash^B_{\text{PLLDG}}$ is relative to the goal base, because the consistency check R_F refers to it.

Definition 3 (PLLDG). *Let $\mathcal{L}$ be a propositional base logic. The language of PLLDG consists of the labelled dyadic goals $G(\alpha|\beta)_L$, with α and β sentences of $\mathcal{L}$, and L a pair (F, p) that consists of a set of sets of strengthenings ($\mathcal{L}$ sentences) together with unlabelled goals (F: fulfillments), and an integer (p: the phase).*

Let ρ be a phasing function that associates with each proof rule below an integer called its phase, let a formula $G(\alpha|\beta)_{(\{\{G(\alpha|\beta)\}\},0)}$, where $\alpha \wedge \beta$ is consistent in $\mathcal{L}$, be called a premise of PLLDG, and let the (extended) goal base $B = (S, I)$ be a set of premises S together with an irreflexive and symmetric relation I on $S \times S$, where $I(g_1, g_2)$ should be read as 'the goal g_1 is independent of the goal g_2.' The goal base B derives a goal $G(\alpha|\beta)_L$ in the phased labelled logic of defeasible conditional goals PLLDG for ρ if $G(\alpha|\beta)_L$ can be derived from S with the inference rules below, extended with the following two conditions $R = R_F + R_p$.

R_F: *$G(\alpha|\beta)_{(F,p)}$ may only be derived from $G(\alpha_1|\beta_1)_{(F_1,p_1)}$ (and $G(\alpha_2|\alpha_2)_{(F_2,p_2)}$) if for each set $F_i \in F$, $G(\alpha|\beta)$ does not violate a goal that depends on some element of F_i, i.e. for which $I(,)$ does not hold. Any rule is applicable if the fact that the labels of the premises are consistent with each goal of the base implies that the label of the consequence is consistent with each goal of the base.*

R_p: *$G(\alpha|\beta)_{(F,p)}$ may only be derived if $p \geq p_1$ (and $p \geq p_2$) for $G(\alpha_1|\beta_1)_{(F_1,p_1)}$ (and $G(\alpha_2|\beta_2)_{(F_2,p_2)}$) it is derived from. A rule is applicable only if no rule of a higher phase has been applied before.*

The inference rules of PLLDG are replacements by logical equivalents and the following four rules.

$$\text{SA}_R : \frac{G(\alpha \mid \beta_1)_{(F,p)}, R}{G(\alpha \mid \beta_1 \wedge \beta_2)_{(F \times \{\{\beta_2\}\}, \rho(\text{SA}))}}$$

$$\text{CTA}_R : \frac{G(\alpha \mid \beta \wedge \gamma)_{(F_1,p_1)}, G(\beta \mid \gamma)_{(F_2,p_2)}, R}{G(\alpha \wedge \beta \mid \gamma)_{(F_1 \times F_2, \rho(\text{CTA}))}}$$

$$\text{WC}_R : \frac{G(\alpha_1 \mid \beta)_{(F,p)}, R}{G(\alpha_1 \vee \alpha_2 \mid \beta)_{(F,\rho(\text{WC}))}}$$

$$\text{OR}_R : \frac{G(\alpha \mid \beta_1)_{(F_1,p_1)}, G(\alpha \mid \beta_2)_{(F_2,p_2)}, R}{G(\alpha \mid \beta_1 \vee \beta_2)_{(F_1 \cup F_2,\rho(\text{OR}))}}$$

where the product is defined by

$$\{S_1, \ldots, S_n\} \times \{T_1, \ldots, T_m\} = \{S_1 \cup T_1, \ldots, S_1 \cup T_m, \ldots, S_n \cup T_m\}$$

Let $B = (\{G(\alpha_1 \mid \beta_1), \ldots, G(\alpha_n \mid \beta_n)\}, I)$ *be an unlabelled goal base. We write* $B \vdash^B_{\text{PLLDG}} G(\alpha \mid \beta)$ *if and only if there is a labelled goal* $G(\alpha \mid \beta)_L$ *that can be derived from the set of goals* $\{G(\alpha_1|\beta_1)_{(\{\{\alpha_1 \wedge \beta_1\}\},0)}, \ldots, G(\alpha_n|\beta_n)_{(\{\{\alpha_n \wedge \beta_n\}\},0)}\}$. *Moreover, we write* $B \mid\!\sim_{\text{PLLDG}} G(\alpha|\beta)$ *if and only if* $B \vdash^B_{\text{PLLDG}} G(\alpha|\beta)$.

Generalizing results from [16] in an obvious way we can show that the logic derives restricted versions of cumulative transitivity (CT) and the conjunction rule for the consequent (AND) below. Moreover, under certain circumstances (for example when all rules are applicable in the same phase) we can equivalently replace CTA by these two rules.

$$\text{CT} : \frac{G(\alpha \mid \beta \wedge \gamma), G(\beta \mid \gamma)}{G(\alpha \mid \gamma)} \qquad \text{AND} : \frac{G(\alpha \mid \gamma), G(\beta \mid \gamma)}{G(\alpha \wedge \beta \mid \gamma)}$$

In Section 4 we prove that we can recapture two previously proposed systems in our framework (because goals depend on themselves, i.e. I is irreflexive):

- The phased labelled logic of conditionals goals PLLG [16] is the PLLDG in which each goal depends only on itself;
- The labelled logic of conditionals goals LLG [15] is the PLLG with phasing function $\rho(x) = 1$.

However, first we use the logic to illustrate the three reasons why reasoning about goals is non-trivial. We focus on the consistency check on the fulfillments.

3 Examples

The first example concerns conflicts. The logic PLLDG can reason about conflicting goals, because we have $G(p \mid \top), G(\neg p \mid \top) \not\mid\!\sim_{\text{PLLDG}} G(p \wedge \neg p \mid \top)$ as well as $G(p \mid \top), G(\neg p \mid \top) \not\mid\!\sim_{\text{PLLDG}} G(q \mid \top)$ if q is not logically implied by p or $\neg p$. In particular, for any phasing function ρ and independence relation I the consistency check blocks the second derivation step in the following counterintuitive derivation. In the derivation below, as well as in following ones, the first blocked step is represented by a dashed line.

$$\dfrac{\dfrac{\dfrac{\dfrac{G(p|\top)_{(\{\{G(p|\top)\}\},0)}}{G(p \vee q|\top)_{(\{\{G(p|\top)\}\},\rho(\text{WC}))}}\ \text{WC}}{G(p \vee q \mid \neg p)_{(\{\{G(p|\top),\neg p\}\},\rho(\text{SA}))}}\ \text{SA (dashed)} \qquad G(\neg p|\top)_{(\{\{G(\neg p|\top)\}\},0)}}{\dfrac{G(q \wedge \neg p|\top)_{(\{\{G(p|\top),\neg p,G(\neg p|\top))\}\},\rho(\text{CTA}))}}{G(q|\top)_{(\{\{G(p|\top),\neg p,G(\neg p|\top)\}\},\rho(\text{WC}))}}\ \text{WC}}\ \text{CTA}$$

The label of the goal potentially derived by SA contains besides the premise $G(p|\top)$ also the strengthening $\neg p$. Consequently SA cannot be applied, because it would result in a goal with contradictory fulfillment p and strengthening $\neg p$. In general, it is obvious that the two premises $G(p|\top)$ and $G(\neg p|\top)$ cannot be combined in *any* single case, because the derived goal would contain the goals with contradictory fulfillments p and $\neg p$.

The second example concerns context sensitivity. Consider the two goals 'I prefer coffee' $G(c|\top)$ and 'I prefer tea if there is no coffee' $G(t|\neg c)$. The logic PLLDG combines strengthening of the antecedent (SA) with weakening of the consequent (WC) without validating the following counterintuitive derivation of 'I prefer coffee or no tea if there is no coffee' $G(c \vee \neg t|\neg c)$ from the first premise.

$$\dfrac{\dfrac{G(c|\top)_{(\{\{G(c|\top)\}\},0)}}{G(c \vee \neg t|\top)_{(\{\{G(c|\top)\}\},\rho(\text{WC}))}}\ \text{WC}}{G(c \vee \neg t|\neg c)_{(\{\{G(c|\top),\neg c\}\},\rho(\text{SA}))}}\ \text{SA (dashed)}$$

The SA rule cannot be applied, because the label of the potentially derived goal contains besides the premise $G(c|\top)$ also the strengthening $\neg c$. In general, we call a goal $G(\alpha|\beta)$ a secondary goal of the primary goal $G(\alpha_1|\beta_1)$ if and only if $\beta \wedge \alpha_1$ is classically inconsistent, and a secondary goal can neither be derived from a primary one, nor can they be combined in a single case.

The third example concerns overriding. Consider the goals 'Sue prefers to marry John over not' $G(j|\top)$ and 'Sue prefers to marry Fred over not' $G(f|\top)$. The PLLDG derivation below shows that the goal 'Sue prefers to marry John and Fred' $G(j \wedge f|\top)$ can defeasibly be derived, because the conjunction rule can be derived from SA and CTA if $\rho(\text{CTA}) \geq \rho(\text{SA})$.

$$\dfrac{\dfrac{G(j|\top)_{(\{\{G(j|\top)\}\},0)}}{G(j|f)_{(\{G(j|\top),f\}\},\rho(\text{SA}))}}\ \text{SA} \qquad G(f|\top)_{(\{\{G(f|\top)\}\},0)}}{G(j \wedge f|\top)_{(\{\{G(j|\top),f,G(f|\top)\}\},\rho(\text{CTA}))}}\ \text{CTA}$$

Now consider the additional goal 'Sue prefers not to be married to both John and Fred' $G(\neg(j \wedge f)|\top)$. If the goals are independent, i.e. I contains at least $(G(j \mid \top), G(\neg(j \wedge f) \mid \top))$ and $(G(f \mid \top), G(\neg(j \wedge f) \mid \top))$, then the derivation above is still valid, because additional independent premises do not influence valid derivations. Otherwise the derivation is not valid, because the SA rule is blocked. The counterintuitive goal cannot be derived, because the fulfillment of $G(j \wedge f|\top)$ violates the additional premise.

$$\dfrac{\dfrac{G(j|\top)_{(\{\{G(j|\top)\}\},0)}}{G(j|f)_{(\{G(j|\top),f\}\},\rho(\text{SA}))}}\ \text{SA (dashed)} \qquad G(f|\top)_{(\{\{G(f|\top)\}\},0)}}{G(j \wedge f|\top)_{(\{\{G(j|\top),f,(f|\top)\}\},\rho(\text{CTA}))}}\ \text{CTA}$$

Note that the fact that the conjunction rule only holds non-monotonically implies that we have to be careful whether we formalize a goal for $p \wedge q$ by the formula $G(p) \wedge G(q)$ or by $G(p \wedge q)$, where we write $G(\alpha)$ for $G(\alpha|\top)$. Although we have $G(p), G(q) \mid\!\sim_{\text{PLLDG}} G(p \wedge q)$, $G(p \wedge q) \mid\!\sim_{\text{PLLDG}} G(p)$ and $G(p \wedge q) \mid\!\sim_{\text{PLLDG}} G(q)$, they derive different consequences when other goals are added to the goal base. For example with only dependent goals we have $G(p), G(q), G(r) \mid\!\sim_{\text{PLLDG}} G(r|\neg(p \wedge q))$ and $G(p \wedge q), G(r) \not\mid\!\sim_{\text{PLLDG}} G(r|\neg(p \wedge q))$, and with dependent goals we also have $G(p), G(q), G(r|\neg p) \not\mid\!\sim_{\text{PLLDG}} G(r|\neg p \wedge \neg q)$ and $G(p \wedge q), G(r|\neg p) \mid\!\sim_{\text{PLLDG}} G(r|\neg p \wedge \neg q)$. Hence, both representations are incomparable, in the sense that in some circumstances the former is more adventurous, and in other circumstances the latter.

Finally, consider dependent goals in $B_1 = (\{G(p), G(q)\}, \emptyset)$ and independent goals in $B_2 = (\{G(p), G(q)\}, \{(G(p), G(q)), (G(q), G(p))\})$. The fulfillment of $G(p \mid \neg q)$ violates $G(q \mid \top)$, and the former goal can therefore only be derived when the two goals are independent. We thus have $B_1 \not\mid\!\sim_{\text{PLLDG}} G(p|\neg q)$ and $B_2 \mid\!\sim_{\text{LLDG}} G(p|\neg q)$. In other words, we have inheritance to sub-ideal subclasses $(\neg q)$ only for independent goals. Due to lack of space we cannot further discuss the independence relation and the procedure to update it when additional goals are added to the goal base.

4 The relation between PLLDG and PLLG

The following proposition shows that PLLDG derives a subset of PLLG introduced in [16].

Proposition 1. *Let PLLG be the PLLDG in which the consistency check R is replaced by the following check $R_{F-\text{PLLG}}$ on the consistency of the fulfillments of all elements of F.*

$R_{F-\text{PLLG}}$: *$G(\alpha|\beta)_{(F,p)}$ may only be derived if the strengthenings together with the fulfillments of goals of each $F_i \in F$ are consistent.*

Note that the related inference relation $\vdash_{\text{PLLG}}$ of PLLG is not relative to the goal base, because $R_{F-\text{PLLG}}$ does not refer to it. If we have $B \mid\!\sim_{\text{PLLDG}} \alpha$, then we have $B \vdash_{\text{PLLG}} \alpha$ (but obviously not necessarily vice versa).

Proof *We prove that if R_F holds, then $R_{F-\text{PLLG}}$ also holds. Assume $R_{F-\text{PLLG}}$ does not hold. This means that the derived goal violates the goal it is derived from. Therefore R_F does not hold, because each goal depends on itself (I is irreflexive).*

The following proposition shows how PLLDG can be defined as PLLG with additional restrictions.

Proposition 2. *Let $B = (\{G(\alpha_1 \mid \beta_1), \ldots, G(\alpha_n \mid \beta_n)\}, I)$ be an unlabelled goal base, and let the materializations of dependent goals of $G(\alpha_i \mid \beta_i)$ of $B = (S, I)$ be* $MDG(G(\alpha_i|\beta_i) \mid B) = \left\{ \{\beta_j \rightarrow \alpha_j\} \;\middle|\; \begin{array}{c} G(\alpha_j|\beta_j) \in S, \\ \textit{not } I(G(\alpha_i|\beta_i), G(\alpha_j|\beta_j)), \\ \alpha_i \wedge \beta_i \wedge \alpha_j \wedge \beta_j \textit{ is consistent} \end{array} \right\}$

We have $B \mid\!\sim_{\text{LLDG}} G(\alpha|\beta)$ *if and only if there is a labelled goal* $G(\alpha|\beta)_L$ *that can be derived from the set of goals*

$$\{G(\alpha_i|\beta_i)_{\{\{G(\alpha_i|\beta_i)\}\}\times MDG(G(\alpha_i|\beta_i)|B)} \mid 1 \leq i \leq n\}$$

in PLLG*, i.e. in* PLLDG *in which the consistency check* R_F *is replaced by the check* $R_{F-\text{PLLG}}$ *of Proposition 1.*

Proof *We prove that* R_F *implies* $R_{F-\text{PLLG}}$ *and vice versa.* $\Rightarrow$ *Assume* $R_{F-\text{PLLG}}$ *does not hold. This means that there is an* $F_i \in F$ *which is inconsistent, and this* F_i *contains a case with an element of MDG, say* $\{G(\alpha_i|\beta_i), \beta_j \to \alpha_j\}$*. This represents exactly that the derived goal violates the goal* $G(\alpha_j|\beta_j)$ *(that depends on* $G(\alpha_i|\beta_i)$*). Consequently* R_F *does not hold.* $\Leftarrow$ *If* $R_{F-\text{PLLG}}$ *holds then, arguing analogously, there is no violated dependent goal. Hence* R_F *holds.*

In the following section we consider surprising consequences of this relation: *modified versions* of the important phasing theorems of PLLG also hold for PLLDG.

5 The one-phase logic LLDG and four-phase logic 4LLDG

We are interested in the following one- and four-phase logic, that represent the two systems discussed in the introduction with respectively a global consistency constraint and a local consistency constraint with phasing.

Definition 4 (LLDG,4LLDG). *The logic* LLDG *is the* PLLDG *with the phasing function* ρ *defined by* $\rho(\text{SA}) = 1$, $\rho(\text{CTA}) = 1$, $\rho(\text{WC}) = 1$, $\rho(\text{OR}) = 1$*. The logic* 4LLDG *is the* PLLDG *with* $\rho(\text{SA}) = 1$, $\rho(\text{CTA}) = 2$, $\rho(\text{WC}) = 3$, $\rho(\text{OR}) = 4$*.*

In Theorem 1 below we show that the phasing does not restrict the set of derivable goals: for each LLDG derivation there is an equivalent 4LLDG derivation. This proof rewriting theorem is analogous to Theorem 1 for LLG and 4LLG in [16]. We first prove two propositions. The first says that for each case the fulfillments of the premises of the case together with the strengthenings imply the fulfillment of the derived goal.

Proposition 3. *For each goal* $G(\alpha|\beta)_{(F,p)}$ *derived in* PLLDG *we have for each* $F_i \in F$ *that the fulfillments of goals in* F_i *together with the strengthenings in* F_i *classically imply* $\alpha \wedge \beta$*.*

Proof *By induction on the structure of the proof tree. The property trivially holds for the premises, and it is easily seen that the proof rules retain the property.*

The second proposition says that we can rewrite two steps of an LLDG proof into a 4LLDG proof.

Proposition 4. *We can replace two subsequent steps of an* LLDG *derivation by an equivalent* 4LLDG *derivation without violating a consistency constraint.*

Proof *In Proposition 3 in [16] this has been proven for* LLG *and* 4LLG. *It can be generalized for* LLDG *and* 4LLDG, *because* PLLDG *derivations can be considered as* PLLG *derivations with modified labels for premises, see Proposition 2. To illustrate this generalization, we repeat one reversal in Figure 1 below. We have to prove that if the first derivation does not violate R_F, then neither does the second derivation. Call the sets of fulfillments of the three premises respectively F_1, F_2 and F_3. The first item of the label of the first derived conditional is*

$$F_1 \times (F_2 \cup F_3) = (F_1 \times F_2) \cup (F_1 \times F_3)$$

and the first item of the label of the second derived conditional is

$$(F_1 \times \{\gamma_1\} \times F_2) \cup (F_1 \times \{\gamma_2\} \times F_3)$$

The two sets contain equivalent formulas, because each element of F_2 classically implies γ_1 and each element of F_3 classically implies γ_2 (Proposition 3). The other proofs are analogous.

$$\dfrac{G(\alpha|\beta \wedge (\gamma_1 \vee \gamma_2)) \qquad \dfrac{G(\beta|\gamma_1) \quad G(\beta|\gamma_2)}{G(\beta|\gamma_1 \vee \gamma_2)}\,\textsc{or}}{G(\alpha \wedge \beta|\gamma_1 \vee \gamma_2)}\,\textsc{cta}$$

$$\dfrac{\dfrac{\dfrac{G(\alpha|\beta \wedge (\gamma_1 \vee \gamma_2))}{G(\alpha|\beta \wedge \gamma_1)}\,\textsc{sa} \qquad G(\beta|\gamma_1)}{G(\alpha \wedge \beta|\gamma_1)}\,\textsc{cta} \qquad \dfrac{\dfrac{G(\alpha|\beta \wedge (\gamma_1 \vee \gamma_2))}{G(\alpha|\beta \wedge \gamma_2)}\,\textsc{sa} \qquad G(\beta|\gamma_2)}{G(\alpha \wedge \beta|\gamma_2)}\,\textsc{cta}}{G(\alpha \wedge \beta|\gamma_1 \vee \gamma_2)}\,\textsc{or}$$

Fig. 1. e.2. Reversing the order of $\textsc{or}_4$ and $\textsc{cta}_2$

From the two previous propositions follows the following generalization of Theorem 1 of [16]. This theorem illustrates the advantage of labelled deductive systems to inductively generalize of the rewriting of two subsequent steps to the rewriting of arbitrary proofs.

Theorem 1 (Equivalence LLDG **and** 4LLDG**).** *Let B be a goal base. We have $B \mathrel{|\!\!\sim}_{\text{LLDG}} G(\alpha|\beta)$ if and only if $B \mathrel{|\!\!\sim}_{\text{4LLDG}} G(\alpha|\beta)$.*

Proof $\Leftarrow$ *Every* 4LLDG *derivation is a* LLDG *derivation.* $\Rightarrow$ *We can take any* LLDG *derivation and construct an equivalent* 4LLDG *derivation, by iteratively replacing two subsequent steps in the wrong order by several steps in the right order, see Proposition 4.*

Moreover, Theorem 2 shows that in 4LLDG the global consistency check on the label can be replaced by a local consistency check on the conjunction of the antecedent and consequent of the goal. This is different from the analogous Theorem 2 for 4LLG in [16], because for non-defeasible goals a consistency check on phase-1 goals is sufficient, whereas for defeasible goals we have to check phase-2 goals too.

Theorem 2. *Let R_{AC} be the following condition.*

R_{AC}**:** *$G(\alpha|\beta)_{(F,p)}$ may only be derived from $G(\alpha_1|\beta_1)$ (and $G(\alpha_2|\alpha_2)$) if it does not violate a dependent goal (of the premise used to derive $G(\alpha_1|\beta_1)$ (and $G(\alpha_2|\alpha_2)$)).*

Consider any potential derivation of 4LLDG*, satisfying the condition R_p but not necessarily R_F. Then the following conditions are equivalent:*

1. *The derivation satisfies R_F everywhere,*
2. *The derivation satisfies condition R_{AC} throughout phase 1 and 2.*

Proof *Through phase 1 and 2, F in the label has a unique element. The conjunction of antecedent and consequent of the derived goal is equivalent to the conjunction of fulfillments of goals of its label and the strengthenings. Hence, in phase 1 and 2 R_F holds iff R_{AC} holds. In phase 3 and 4 no goals can be violated.*

This proves the completeness of the phased proof theory given in the introduction, because we can simply define atomic goals as phase-1 goals, compound goals as phase-2 goals, and combined goals as phase-4 goals. The goal $G(\alpha|\beta)$ is *an atomic goal* of goal base B if there is a goal $G(\alpha\,|\,\beta') \in B$ such that β classically implies β', i.e. $\beta \vdash \beta'$, and $\alpha \wedge \beta$ does not violate a goal dependent on $G(\alpha|\beta')$. The goal $G(\alpha_1 \wedge \ldots \wedge \alpha_n|\beta)$ is a compound goal of B if there is a set of atomic goals $\{G(\alpha_1|\beta), G(\alpha_2|\beta \wedge \alpha_1), \ldots, G(\alpha_n|\beta \wedge \alpha_1 \wedge \ldots \wedge \alpha_{n-1})\}$ of B such that $\beta \wedge \alpha_1 \wedge \ldots \wedge \alpha_i$ does not violate a goal dependent on the goal used to derive $G(\alpha_{i-1}|\beta \wedge \alpha_1 \wedge \ldots \wedge \alpha_{i-2})$ for $i = 2 \ldots n$. The goal $G(\alpha|\beta_1 \vee \ldots \vee \beta_n)$ is a combined goal of B if there is a set of compound goals $\{G(\alpha_1|\beta_1), \ldots, (G(\alpha_n|\beta_n)\}$ of B such that $\alpha_i \vdash \alpha$ for $i = 1 \ldots n$.

6 Conclusions

In this paper we studied a labelled logic of defeasible goals, with several desirable properties. The logic does not have the identity or conflict averse strategies, but it has restricted strengthening of the antecedent, weakening of the consequent, the conjunction rule, cumulative transitivity and the disjunction rule. We introduced phasing in the logic and we proved two phasing theorems. Phasing makes the proof theory more efficient, because only a single order of rule application has to be considered.

Restricted applicability is of course well-known from logics of defeasible reasoning, where monotony and transitivity only hold in a restricted sense and where restricted alternatives of monotony have been developed, e.g. cautious and rational monotony. However, it is important to note that the logic of defaults is quite different from the logic of defeasible goals. For example, according to the generally accepted Kraus-Lehmann-Magidor paradigm [7] the identity (called supra-classicallity) holds in the logic of defaults. Moreover, the conjunction rule holds without restrictions, whereas in the logic of goals it has to be restricted to formalize the marriage example. This distinction is not surprising, because the quantitative counterparts – utilities and probabilities – also have completely different properties. A systematic study of these two types of non-monotonic reasoning is the subject of present investigations.

References

1. F. Bacchus and A.J. Grove. Utility independence in a qualitative decision theory. In *Proceedings of KR'96*, pages 542–552, 1996.
2. C. Boutilier. Toward a logic for qualitative decision theory. In *Proceedings of the KR'94*, pages 75–86, 1994.
3. J. Doyle, Y. Shoham, and M.P. Wellman. The logic of relative desires. In *Sixth International Symposium on Methodologies for Intelligent Systems*, Charlotte, North Carolina, 1991.
4. J. Doyle and M.P. Wellman. Preferential semantics for goals. In *Proceedings of the AAAI'91*, pages 698–703, Anaheim, 1991.
5. D. Gabbay. *Labelled Deductive Systems*, volume 1. Oxford University Press, 1996.
6. Z. Huang and J. Bell. Dynamic goal hierarchies. In *Intelligent Agent Systems: Theoretical and Practical Issues*, LNAI 1027, pages 88–103. Springer, 1997.
7. S. Kraus, D. Lehmann, and M. Magidor. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, 44:167–207, 1990.
8. J. Lang. Conditional desires and utilities - an alternative approach to qualitative decision theory. In *Proceedings of the ECAI'96*, pages 318–322, 1996.
9. D. Makinson. On a fundamental problem of deontic logic. In P. McNamara and H. Prakken, editors, *Norms, Logics and Information Systems. New Studies on Deontic Logic and Computer Science*. IOS Press, 1998.
10. D. Makinson and L. van der Torre. The logic of reusable propositional output. *Submitted*, 1999.
11. J.D. Mullen. Does the logic of preference rest on a mistake? *Metaphilosophy*, 10:247–255, 1979.
12. J. Pearl. From conditional oughts to qualitative decision theory. In *Proceedings of the UAI'93*, pages 12–20, 1993.
13. S.-W. Tan and J. Pearl. Qualitative decision theory. In *Proceedings of the AAAI'94*, 1994.
14. S.-W. Tan and J. Pearl. Specification and evaluation of preferences under uncertainty. In *Proceedings of the KR'94*, pages 530–539, 1994.
15. L. van der Torre. Labeled logics of conditional goals. In *Proceedings of the ECAI'98*, pages 368–369. John Wiley & Sons, 1998.
16. L. van der Torre. Phased labeled logics of conditional goals. In *Logics in Artificial Intelligence*, LNAI 1489, pages 92–106. Springer, 1998.
17. L. van der Torre and Y. Tan. Cancelling and overshadowing: two types of defeasibility in defeasible deontic logic. In *Proceedings of the IJCAI'95*, pages 1525–1532. Morgan Kaufman, 1995.