On Default Representation of Defeasible Inference and Specificity*

Alexander Bochman
Computer Science Department,
Holon Institute of Technology, Israel
bochmana@hit.ac.il

Abstract

We suggest a new representation of defeasible entailment and specificity in the framework of default logic. The representation is based on augmenting the underlying classical language with the language of conditionals having its own (monotonic) internal logic. It is shown, in particular, that inheritance reasoning can be naturally represented in this framework, and generalized to the full classical language.

Introduction

The problem of nonmonotonic, defeasible inference can be seen as the main objective, as well as the main problem of the general theory of nonmonotonic reasoning. An impressive success has been achieved in our understanding of it, and mainly in realizing how many different forms and aspects it has. Many formalisms have been developed and implemented that capture significant aspects of nonmonotonic inference, though a unified picture has not yet been emerged. In this study we will suggest a new representation of defeasible inference in default logic that combines, in effect, the insights and advantages of a number of previous approaches to this problem, hopefully without inheriting their shortcomings.

Suppose we have a default rule $A \to B$ saying "A normally implies B". On a most natural, commonsense understanding, such a rule represents a claim that A implies B, given some additional (unmentioned and even not fully known) assumptions that are presumed to hold in normal circumstances. Thus, a default causal rule $TurnKey \to CarStarts$ states that if I turn the key, the car will start given the normal conditions such as there is a fuel in the tank, the battery is ok, etc. etc. An important aspect of our understanding of such default rules is that the default assumptions required for different rules, say $A \to B$ and $A \to C$, are usually presumed to be independent. Consequently, a violation of $A \to B$ does not imply rejection of $A \to C$. Note that already this presumption is incompatible with the usual assumptions about normality made in preferential inference (see (**?**)): according to the latter, a violation of $A \to B$ means that the situation at hand is abnormal with respect to A , so we are not entitled to infer anything that normally follows from A. Consequently, default reasoning is not directly captured by the preferential approach.

It was John McCarthy (see (**?**; **?**)) who has suggested to represent default rules $A \to B$ as classical implications of the form $A \wedge \neg ab \supset B$, where ab is a new 'abnormality' proposition that accumulates the conditions for rejection of the source rule. In fact, viewed as a formalism for nonmonotonic reasoning, the central concept of McCarthy's circumscriptive method is not circumscription itself, but his notion of an *abnormality theory* - a set of classical conditionals containing the abnormality predicate ab that provides a representation for default information.

McCarthy's representation can be seen as a particular instantiation of our description above, where $\neg ab$ serves as an abstract representation of the default assumptions required for the inference in question. The default character of these assumptions was captured in McCarthy's theory by a circumscription policy that minimized abnormality (and thereby maximized the acceptance of the corresponding normality claims $\neg ab$). Since then, this representation of default rules using auxiliary (ab)normality propositions has been employed both in applications of circumscription, and in many other theories of nonmonotonic inference in AI, sometimes in alternative logical frameworks. Some major examples are inheritance theories (**?**), logic-based diagnosis (**?**), general representation of defaults in (**?**) and reasoning about time and action. Note also that naming of defaults (as in (**?**)) can also be viewed as a species of this idea. Finally, the approach described below can also be seen as a development of this representation.

Abnormality theories have brought out, however, several problems in the application of circumscription to commonsense reasoning. One of the most pressing was the specificity problem arising when there are conflicting defaults. In combining two defaults, "Birds fly" and "Penguins can't fly", the specificity principle naturally suggests that the second, more specific, default should be preferred. A general approach to handle this problem in circumscription, suggested in (**?**) and endorsed in (**?**), was to impose priorities among minimized predicates and abnormalities. The corresponding variant of circumscription has been called prioritized circumscription.

In reading McCarthy's papers on circumscription, one

*Dedicated to John McCarthy on his 80th birthday.

cannot help feeling uneasiness with which he adopts the abnormality predicates into the language, since he thought that this compels us to introduce abnormalities as new entities into our ontology (the things that exist). Moreover, in its developed form, described in (**?**), the representation required a relativization of abnormality claims with respect to particular *aspects*, so that some aspects can be abnormal without affecting others. This modification was required in order to cope with the above mentioned presumption about independence of default assertions with the same antecedent and different conclusions. As was noted by McCarthy, the aspects themselves are abstract entities, and their unintuitiveness is somewhat a blemish on the theory.

Our approach below is based on the idea that the above (ab)normality propositions can be assigned a natural meaning that would also facilitate their conscious and coherent use in nonmonotonic reasoning. More specifically, we argue that the default assumptions of a defeasible rule $A \rightarrow B$ provide a link (an information channel) that sanctions the inference from A to B. In other words, they jointly function as a conditional, that we will denote by A/B, that, once accepted, allows us to infer B from A. Accordingly, we will slightly 'unfold' the normality assumption $\neg ab$, and represent $A \rightarrow B$ as the classical implication $A \wedge (A/B) \supset B$. The default character of this inference will be captured by requiring that A normally implies A/B, that is, $A{\rightarrow}(A/B)$[1]. It will be shown below, however, that the latter rule can be represented simply as Reiter's normal default.

It is important to note that the change we have made so far to McCarthy's representation is purely terminological. This means, in particular, that the implementation of the theory of defeasible inference that will be described in the sequel can be made, in principle, by switching back to abnormality predicates and established formalisms of dealing with them. The new representation naturally suggests, however, that the (ab)normality claims should have its own internal logic. In fact, we will stipulate below that the conditionals A/B should satisfy at least the usual rules of supraclassical Tarski consequence relations. It is this internal logic that will allow us to formulate purely logical principles that will govern a proper interaction of defeasible rules in cases of conflict.

The Language and Logic of Conditionals

Our basic language L_0 will be a classical propositional language with the usual connectives and constants $\{\wedge, \vee, \neg, \supset, \mathbf{t}, \mathbf{f}\}$. $\vDash$ will stand for the classical entailment, while Th will denote the associated provability operator.

As a first step, we will augment the language L_0 by adding new propositional atoms of the form A/B, where A and B are classical propositions of L_0. The conditionals A/B will be viewed as propositional atoms of a new type, so nesting of conditionals will not be allowed. Still, the new propositional atoms will be freely combined with ordinary ones using the classical propositional connectives. We will denote the resulting language by L_c.

[1]A similar representation lies at the basis of the approach developed in (**?**).

The essence and main functional role of our conditionals will be expressed by adopting the following axiom:

MP $\quad (A \wedge A/B) \supset B.$

In addition, conditionals will be viewed as ordinary inference rules that are 'reified' in the object language. Accordingly, we will require them to satisfy the inference rules of a supraclassical Tarski consequence relation. It can be shown that the following postulates are sufficient for this purpose:

$$\text{If } A \vDash B, \text{ then } A/B. \qquad \text{(Dominance)}$$

$$\frac{A/B \quad B/C}{A/C} \qquad \text{(Transitivity)}$$

$$\frac{A/B \quad A/C}{A/(B \wedge C)} \qquad \text{(And)}$$

Remark. The above conditional logic can be given a complete semantic interpretation. In fact, our conditional logic coincides with the logic of *monotonic* consequence relations, sketched in (**?**), and the semantics of simple preferential models, described in that paper (namely, models with an empty preference relation) has been shown to be adequate for the latter. The semantic representation of our language will play no role in our subsequent constructions, however, so we will omit its detailed description.

It should be noted that the above logic describes the logical properties of arbitrary conditionals, not only default ones. The difference between the two will be reflected in the representation of defeasible conditionals in the framework of default logic, described next.

Defeasible Inference in Default Logic

We will describe now a modular representation of defeasible rules $A \rightarrow B$ in Reiter's default logic (**?**). Due to space limitations, we will refrain from a detailed description of default logic, but only fix the notation.

A *default theory* is a pair (W, D), where W is a set classical propositions (the axioms), and D is a set of default rules of the form $A : b \vdash C$, where A, C are propositions and b a finite set of propositions. A is called a *prerequisite* of the rule, b a set of its *justifications*, and C - its *conclusion*. The notion of an extension of a default theory is defined as usual: for a set s of propositions, we define $\mathcal{D}(s)$ as the set of all propositions that are derivable from W using the classical entailment and the following inference rules:

$$\{A \vdash C \mid A : b \vdash C \in D \;\&\; \neg B \notin s, \text{ for any } B \in b\}.$$

Then s is an *extension* of the default theory iff $s = \mathcal{D}(s)$.[2]

For the present case, we will suppose that our default theory is defined in the conditional language L_c and respects the corresponding logic of conditionals, that is, W includes MP and Dominance axioms[3], while D includes Transitivity and And as strict rules (without justifications).

Now, for each defeasible rule $A{\rightarrow}B$, we accept the normal default

$$A : A/B \vdash A/B.$$

[2]Cf. (**?**) for a similar definition.

[3]Namely, W should include all A/B such that $A \vDash B$.

Finally, we have two natural options for representing strict (non-defeasible) rules in this framework. A more cautious understanding would lead to representing a strict rule $A \Rightarrow B$ as a strict default rule $A : \vdash B$ without justifications. A more 'classical' understanding would amount to representing $A \Rightarrow B$ as a material implication $A \supset B$; this could be achieved in our framework simply by adding A/B to the set W of axioms.

Specificity and Commitment

The default theory described so far is still insufficient, of course, for capturing defeasible inference, since it does not take into account the *principle of specificity*. In fact, it can be shown that such a default theory is practically indistinguishable from the default theory obtained by representing every defeasible rule $A \to B$ as a normal default rule $A : B \vdash B$.

Fortunately, we now have sufficient means for expressing the specificity principle in a simple and transparent way. The formulation of this principle, given below, can be seen as the main contribution of this study.

Taken literally, the specificity principle states that more specific default rules should override less specific rules in the case of conflict. In a simplest case, this pertains to the conflict between the rules $A \to C$ and $A \wedge B \to \neg C$, in which case the second rule should override the first. There are, however, less direct cases in which one of the conflicting rules is deemed more specific than the other, and, as we all know, the literature is abundant with the attempts to define a more general notion of specificity.

We claim that the specificity principle is a consequence of a more general principle that we will call the *principle of commitment*. According to the latter, by asserting a defeasible rule "If A, then normally B", we are also committing ourselves to the claim that no combination of accepted rules could allow us to infer $\neg B$ from A. A clear expression of this principle can be found already in (**?**):

> "if "p's are q's" is a default and if we know p(c), then all of the objections that could be raised about q(c) that follow from p(c) have already been taken into account when building the knowledge base."

The principle of commitment is related also to the principle of *direct inference* stated in (**?**). According to the latter, if $A \to B$ is accepted, then A should always imply B in the case A is the only evidence we have.

For our present purposes, a special case of the commitment principle can be expressed as follows: If $A \to B$ is a default rule, and A is known to hold, then any conditional of the form A/C can be accepted only if we reject the conditional $C/\neg B$ (otherwise we would have a counterargument against $A \to B$).

Commitment If $A \to B$ is accepted, then, for any proposition C,

$$A, A/C : \vdash \neg(C/\neg B).$$

As an important special case, the commitment principle implies that acceptance of $A \to B$ compels us to reject the opposite conditional (since A/A always holds):

$$A : \vdash \neg(A/\neg B).$$

As a result, if we have both $A \to B$ and $A \to \neg B$ in the default theory, each of the corresponding conditionals A/B and $A/\neg B$ will be 'disabled', so no conclusion will be derived from them.

The following couple of examples will show the impact of accepting the above commitment principle.

Example 1. Consider a generalized Penguin-Bird theory $\{P \to B, B \to F, P \to \neg F\}$. As could be expected, given the fact B, the corresponding default theory has a unique extension that contains B and F. Now, given the fact P instead, the resulting default theory also has a unique extension that includes this time P, B and $\neg F$. The spurious extension containing F is blocked due to the commitment principle for $P \to \neg F$ that implies $P,\ P/B : \vdash \neg(B/F)$. Note, however, that the situation is not symmetric, since the commitment to $B \to F$ does not allow us to reject $P \to \neg F$.

The following example from (**?**) shows that the above representation deals correctly with the interplay of specificity and evidence, unlike the representations such as prioritized circumscription or Geffner's conditional entailment that are based on establishing context-independent priorities among default rules.

Example 2. Let us consider the following default theory $\{A \to M,\ S \to \neg M,\ S \to Y,\ Y \Rightarrow A\}$ that represents, respectively, default assertions that adults are normally married, students are normally not married, students are normally young adults, and the strict rule "young adults are adults". For the evidence S, the corresponding default theory will have a single extension containing $\neg M$. This extension will contain also S/Y and Y/A, so it will include S/A (by Transitivity). As a result, the default conditional A/M will be rejected in this case due to the commitment to $S \to \neg M$. Note, however, that the 'priority' of $S \to \neg M$ over $A \to M$ is not absolute here, since it depends on the acceptance of S/A. Accordingly, given a more specific evidence $S \wedge \neg Y \wedge A$, the conditional S/A will no longer be acceptable (since the default S/Y is refuted by MP), so the resulting default theory will have two extensions, one containing M, another containing $\neg M$. In other words, as should be expected, the marital status of non-young students cannot be decided.

As a matter of fact, the commitment principle constitutes a generalization of the specificity rules stated in (**?**) and especially in (**?**) (the latter being formulated, however, in the framework of logic programming). As a result, we have an opportunity to provide a straightforward representation of defeasible inheritance in our framework.

Defeasible Inheritance

Defeasible inheritance nets is a logical framework that has been originally designed to capture reasoning in taxonomic hierarchies that allowed to have exceptions. The theory of reasoning in such taxonomies has been called *nonmonotonic inheritance* (see (**?**) for an overview). The guiding principle in resolving potential conflicts in such hierarchies was a *specificity principle* ((**?**; **?**)): more specific information should override more generic information in cases of conflict. Though obviously related to nonmonotonic reasoning, nonmonotonic inheritance relied more heavily on graph-based representations than on traditional logical tools. Nevertheless, it has managed to provide a plausible analysis of reasoning in this restricted context.

Let Γ be a consistent defeasible inheritance network. A credulous extension of Γ is defined as usual (see the Appendix), with the only simplification that it is restricted to the set of paths from object nodes (as in (**?**)).

For a propositional atom q, $\hat{q}$ will denote a corresponding literal, that is either q, or $\neg q$, while $\neg\hat{q}$ will denote the literal complementary to $\hat{q}$.

$D(\Gamma)$ will denote the default theory corresponding to Γ as follows[4]:

- any object link will correspond to an axiom $\hat{p} \in W$;
- every defeasible link will correspond to a defeasible rule of the form $p \to \hat{q}$, so $p : p/\hat{q} \vdash p/\hat{q}$ will be added to the default rules, as well as the corresponding Commitment rule:

$$p,\, p/C : \vdash \neg(C/\neg\hat{q}).$$

Then the following theorem shows, in effect, that the resulting default theory provides an exact formalization of defeasible nets.

Theorem. *A set of paths Φ is a credulous extension of Γ if and only if there is an extension u of $D(\Gamma)$ such that Φ coincides with the set of paths constructed from the set of links $\{p \to \hat{q} \in \Gamma \mid p/\hat{q} \in u\}$.*

The gain in simplicity and modularity provided by the above representation could be made vivid by comparing it with the much more complex translation of defeasible inheritance into default logic described in (**?**).

Conclusions

Nonmonotonic reasoning is not just a syntax plus nonmonotonic semantics. An account of the underlying *logic* behind our commonsense reasoning can provide an immense improvement in the quality of representations. In our case, it has been shown that when the normality assumptions mediating defeasible rules are represented as conditionals having a relatively simple underlying logic, the resulting representation has allowed us to capture defeasible inheritance and specificity, generalized to the full classical language. Furthermore, all that was required to achieve a proper nonmonotonicity in this logical setting was the basic formalism of Reiter's normal defaults.

References

Dung, P. M., and Son, T. C. 2001. An argument-based approach to reasoning with specificity. *Artificial Intelligence* 133:35–85.

Etherington, D., and Reiter, R. 1983. On inheritance hierarchies with exceptions. In *Proceedings of AAAI-83*, 104–108.

Geffner, H. 1992. *Default Reasoning. Causal and Conditional Theories*. MIT Press.

Gelfond, M.; Lifschitz, V.; Przymusińska, H.; and Truszczyński, M. 1991. Disjunctive defaults. In *Proc. Second Int. Conf. on Principles of Knowledge Representation and Reasoning, KR'91*, 230–237.

Horty, J. F. 1994. Some direct theories of nonmonotonic inheritance. In Gabbay, D. M.; Hogger, C. J.; and Robinson, J. A., eds., *Handbook of Logic in Artificial Intelligence and Logic Programming* **3***: Nonmonotonic Reasoning and Uncertain Reasoning*. Oxford: Oxford University Press.

Konolige, K., and Myers, K. L. 1989. Representing defaults with epistemic concepts. *Computational Intelligence* 5:32–44.

Kraus, S.; Lehmann, D.; and Magidor, M. 1990. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence* 44:167–207.

Lifschitz, V. 1985. Computing circumscription. In *Proc. 9th Int. Joint Conf. on Artificial Intelligence, IJCAI-85*, 121–127. Morgan Kaufmann.

McCarthy, J. 1980. Circumscription – a form of nonmonotonic reasoning. *Artificial Intelligence* 13:27–39.

McCarthy, J. 1986. Applications of circumscription to formalizing common sense knowledge. *Artificial Intelligence* 13:27–39.

Poole, D. 1985. On the comparison of theories: Preferring the most specific explanation. In *Proceedings Ninth International Joint Conference on Artificial Intelligence*, 144–147.

Poole, D. 1988. A logical framework for default reasoning. *Artificial Intelligence* 36:27–47.

Poole, D. 1990. Dialectics and specificity: conditioning in logic-based hypothetical reasoning (preliminary report). In *Proc. 8th biennial conf. of the Canadian Society for Computational Studies of Intelligence*, 69–76. San Francisco, CA: Morgan Kaufmann.

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

Reiter, R. 1987. A theory of diagnosis from first principles. *Artificial Intelligence* 32:57–95.

Touretzky, D. S. 1986. *The Mathematics of of Inheritance Systems*. Los Altos: Morgan Kaufmann.

You, J.-H.; Wang, X.; and Yuan, L. Y. 1999. Compiling defeasible inheritance networks to general logic programs. *Artificial Intelligence* 113:247–268.

Appendix. Proof of the Theorem

A defeasible inheritance network Γ is a tuple (N, E) where N is a set of nodes and E a set of positive and negative links between nodes. Nodes are divided into two disjoint classes: object nodes, and property nodes. An object node can only be used as a root node. A link is called an object link if its root is an object node. A *path* is a sequence of links such that the head of a preceding link coincides with the root of the next link in the sequence, and all the links in the sequence,

[4] In order to simplify the representation, we will only consider defeasible networks without strict links.

except possibly the last, are positive. A path is *positive*, if all its links are positive, otherwise it is a *negative* path.

Path constructibility. Suppose that Φ is a path set of Γ. A path σ is *constructible* in Φ iff (i) it is an object link, or (ii) σ consists of a prefix $\tau \in \Phi$ and the last link that belongs to E.

Conflict. A positive (res., negative) path σ is *conflicting* in Φ iff $\sigma \in \Phi$ and Φ contains a negative (resp., positive) path with the same beginning and end nodes.

Off-path preemption. Defined as usual.

A path σ is *defeasibly inheritable* in Φ iff it is constructible, not conflicting and not preempted in Φ.

Definition. A set Φ of paths is a *credulous extension* of a net Γ if it coincides with the set of paths that are defeasibly inheritable in Φ.

Theorem. *A set of paths Φ is a credulous extension of Γ if and only if there is an extension u of $D(\Gamma)$ such that Φ coincides with the set of paths constructed from the set of links $\{p \to \hat{q} \in \Gamma \mid p/\hat{q} \in u\}$.*

Proof. ($\Rightarrow$) If Φ is a credulous extension of Γ, let $l(\Phi)$ be the set of all non-object links appearing on paths of Φ. Also, let $R = \{p/\hat{q} \mid p \to \hat{q} \in l(\Phi)\}$, and define $Cl(R)$ to be the set of all conditionals A/B that are derivable from R by the rules Dominance, Transitivity and And. It can be easily shown that $p/\hat{q} \in Cl(R)$ if and only if Φ contains a path via p to $\hat{q}$ (cf. Lemma C3 in (**?**)).

Let u be the closure of the set $W \cup R$ with respect to the classical entailment and the strict rules of $D(\Gamma)$, i.e., Transitivity, And and Commitment.

Lemma 1. *If q is a propositional atom, then $\hat{q} \in u$ iff Φ supports $\hat{q}$.*

Proof. If $\hat{q}$ is supported by Φ, there is a path σ in Φ from an object link to $\hat{q}$. Since $p_i/p_{i+1} \in u$, for any link $p_i \to p_{i+1}$ on this path, we have $p_i \supset p_{i+1} \in u$ (by MP). Clearly then $\hat{q} \in u$. To prove the other direction, we define a propositional theory $v = \mathrm{Th}(v_s \cup Cl(R) \cup N(R))$, where v_s is the set of ordinary literals that are supported in Φ, and $N(R) = \{\neg(A/B) \mid A/B \notin Cl(R)\}$. Note that, for any conditional atom A/B, we have either $A/B \in v$, or $\neg(A/B) \in v$. We will show that $u \subseteq v$.

We will demonstrate first that if $A/B \in Cl(R)$, then $v_s \vDash A \supset B$, by induction on the derivations of A/B. If $p/\hat{q} \in R$, then Φ contains a path that includes $p \to \hat{q}$, so $\hat{q}$ is supported, and consequently $v_s \vDash p \supset \hat{q}$. Now, if $A \vDash B$, then clearly $v_s \vDash A \supset B$. If $A/(B \wedge C)$ has been obtained from A/B and A/C by the rule And, then $v_s \vDash A \supset B$ and $v_s \vDash A \supset C$ by the inductive assumption, so $v_s \vDash A \supset B \wedge C$. The proof is similar if A/C has been obtained from A/B and B/C by Transitivity. Hence the claim holds.

Now we can show that the axiom MP belongs to v. If $A/B \notin v$, then $\neg(A/B) \in v$, so $A \wedge (A/B) \supset B \in v$. Assume that $A/B \in v$. Then by the preceding claim $v_s \vDash A \supset B$, and hence again $A \wedge (A/B) \supset B \in v$. Consequently, all the axioms of $D(\Gamma)$ are included in v. In addition, v is closed with respect to Transitivity and And (since it includes $Cl(R)$). Finally, we will show that v is closed wrt the commitment rules. Suppose that $p \to \hat{q} \in \Gamma$, $p \in v$, $p/A \in v$, but $\neg(A/\neg\hat{q}) \notin v$. Then $A/\neg\hat{q} \in v$, and consequently $p/\neg\hat{q} \in v$ by Transitivity. The latter implies $p/\neg\hat{q} \in Cl(R)$, and hence Φ should contain a path σ from an object link via p to $\neg\hat{q}$, which is impossible, since it is preempted by $p \to \hat{q}$. Thus, v includes W and is closed wrt all the strict rules of $D(\Gamma)$. Consequently, $u \subseteq v$.

Assume now that $\hat{q}$ is not supported in Φ. Clearly, v_s is exactly the set of ordinary literals in v, so $\hat{q} \notin v$, and therefore $\hat{q} \notin u$. This completes the proof of the lemma. $\square$

We will show now that u is an extension of $D(\Gamma)$, that is, $u = \mathcal{D}(u)$.

For the inclusion $u \subseteq \mathcal{D}(u)$, we show first that if $p \to \hat{q} \in l(\Phi)$, then $p \wedge (p/\hat{q}) \in \mathcal{D}(u)$. Now, $p \to \hat{q} \in l(\Phi)$ only if Φ contains a path σ that supports p. We will prove the claim by induction on the length of σ. If $p \to \hat{q}$ is the first non-object link, then $p \in W$. In addition, we can apply $p : p/\hat{q} \vdash p/\hat{q}$ to derive $p/\hat{q}$ (since $p/\hat{q} \in u$). Assume now that σ is a path of length n having $r \to p$ as the last link. By the inductive assumption, $r \wedge (r/p) \in \mathcal{D}(u)$, so $p \in \mathcal{D}(u)$ by MP: $r \wedge (r/p) \supset p$. Consequently, $p/\hat{q} \in \mathcal{D}(u)$ by the rule $p : p/\hat{q} \vdash p/\hat{q}$, and we are done.

The above claim implies $\{p/\hat{q} \mid p \to \hat{q} \in l(\Phi)\} \subseteq \mathcal{D}(u)$, so by the definition of u we immediately conclude $u \subseteq \mathcal{D}(u)$.

For the inclusion $\mathcal{D}(u) \subseteq u$, it is sufficient to show that u is closed with respect to all the rules of the default theory $D(\Gamma)$. First, u includes W. Suppose that $p \to \hat{q} \in \Gamma$ but the rule $p : p/\hat{q} \vdash p/\hat{q}$ does not hold in u, that is $p \in u$, $p/\hat{q} \notin u$ and $\neg(p/\hat{q}) \notin u$. Now, $p/\hat{q} \notin u$ implies $p \to \hat{q} \notin l(\Phi)$, which can happen only if it is either conflicted or preempted in Φ. Suppose first that $p \to \hat{q}$ is conflicted in Φ. Then $\neg\hat{q}$ is supported by Φ, and hence $\neg\hat{q} \in u$. Consequently $\neg(p/\hat{q}) \in u$ by MP, contrary to our assumptions. Suppose now that $p \to \hat{q}$ is preempted in Φ. Then Γ contains a link $r \to \neg\hat{q}$ such that there is a positive path σ via r to p in Φ. Let us consider the sub-path of σ from r to p. Since $p_i/p_{i+1} \in u$, for any link $p_i \to p_{i+1}$ that belongs to this sub-path, we obtain $r/p \in u$ by transitivity. Then by the commitment rule $\neg(p/\hat{q}) \in u$ contrary to our assumptions. The obtained contradiction shows that u is closed with respect to all the rules of $D(\Gamma)$, and consequently $\mathcal{D}(u) \subseteq u$ holds.

Finally, we have to show that $l(\Phi)$ coincides with $\{p \to \hat{q} \in \Gamma \mid p/\hat{q} \in u\}$. Assume that $p \to \hat{q} \in \Gamma$ and $p/\hat{q} \in u$. By the construction of u, this can happen only if $p/\hat{q} \in Cl(R)$, and therefore there is a path in Φ of the form $\sigma p \tau \hat{q}$. Let us consider the path $\sigma_1 = \sigma p \to \hat{q}$. This path is constructible in Φ, and it is clearly neither conflicted, nor preempted in Φ. Since Φ is a credulous extension, we conclude $\sigma_1 \in \Phi$, and therefore $p \to \hat{q} \in l(\Phi)$.

($\Leftarrow$) Suppose that u is an extension of $D(\Gamma)$, and define Φ to be the set of paths constructed from the object links and the links $\{p \to \hat{q} \in \Gamma \mid p/\hat{q} \in u\}$. Let u_Φ denote the set of all literals that are supported by Φ. If $\hat{p} \in u_\Phi$, there is a path $\sigma \in \Phi$ from an object link to $\hat{p}$. Since $p_i/p_{i+1} \in u$, for any link $p_i \to p_{i+1}$ on this path, we have $p_i \supset p_{i+1} \in u$ (by MP). Clearly then $\hat{p} \in u$. In addition, we will need the following

Lemma 2. *$A \in u$ only if $u_\Phi \models A$, and $A/B \in u$ only if $u_\Phi \models A \supset B$.*

Proof. Since $u = \mathcal{D}(u)$, we will prove these two claims by a simultaneous induction on the derivations in $\mathcal{D}(u)$.

If $\hat{p} \in W$ since $\hat{p}$ corresponds to an object link, then clearly $\hat{p} \in u_\Phi$. Also, if $A/B \in W$ by Dominance, then $A \models B$, and hence $u_\Phi \models A \supset B$.

If $p/\hat{q}$ has been obtained by the default rule $p : p/\hat{q} \vdash p/\hat{q}$, then $p \to \hat{q} \in \Phi$ and $p \in u$. By the inductive assumption, $u_\Phi \models p$, so $p \in u_\Phi$, and hence p is supported by Φ. Therefore, $\hat{q}$ is also supported by Φ, and hence $u_\Phi \models p \supset \hat{q}$.

If $A/(B \wedge C)$ has been obtained from A/B and A/C by the rule And, then $u_\Phi \models A \supset B$ and $u_\Phi \models A \supset C$ by the inductive assumption, so $u_\Phi \models A \supset B \wedge C$. The proof is similar if A/C has been obtained from A/B and B/C by Transitivity.

Finally, the axiom $A \wedge A/B \supset B$ cannot be used for deriving new conditionals, but only for deriving $A \supset B$ when A/B has been proved. But in this case $u_\Phi \models A \supset B$ already by the inductive assumption, and we are done. □

Now we will show that Φ is a credulous extension of Γ. We will prove first that Φ is defeasibly inheritable in Φ, that is, any path in Φ is constructible, conflict-free and not preempted in Φ. Now, any path in Φ is constructible by the definition. If there is a conflicted path in Φ, then there is an atom p, such that both p and $\neg p$ belong to u, which is impossible. Finally, assume that σ is a preempted path in Φ, and $p \to \hat{q}$ is the last link on σ. Then there exist a link $r \to \neg\hat{q}$ and a path τ via r to p that belongs to Φ. Now, $p/\hat{q} \in u$, and $p_i/p_{i+1} \in u$, for any link $p_i \to p_{i+1}$ on the sub-path of τ from r to p. Therefore, $r/p \in u$ and hence by commitment $\neg(p/\hat{q}) \in u$ - a contradiction. Therefore, σ is not preempted in Φ.

Finally, we will show that any path that is defeasibly inheritable in Φ also belongs to Φ. Suppose that σ is defeasibly inheritable in Φ. If it is an object link, it is in Φ. Assume that σ is composed of a prefix $\tau \in \Phi$ and the last link $p \to \hat{q}$. Then p is supported in Φ, and therefore $p \in u$. Next we are going to show that $\neg(p/\hat{q}) \notin u$. Suppose that $\neg(p/\hat{q}) \in u$. Since $u = \mathcal{D}(u)$, we have that $\neg(p/\hat{q})$ should be derivable from W using the strict and active normal default rules of $D(\Gamma)$. As can be seen, this can happen only if $\neg(p/\hat{q})$ is obtained either (i) from the axiom MP when $p \wedge \neg\hat{q} \in u$, or (ii) by the commitment rule $r,\ r/p : \vdash \neg(p/\hat{q})$, given a link $r \to \neg q \in \Gamma$ and the fact $r/p \in u$. In the case (i) we have that $\neg\hat{q}$ is supported by Φ, and therefore σ is conflicted in Φ, contrary to our assumptions. In the case (ii) we have that there should be a path in Φ via r to p, and therefore σ is preempted in Φ, which is impossible. Thus, $\neg(p/\hat{q}) \notin u$, and hence we can apply the default rule $p : p/\hat{q} \vdash p/\hat{q}$ and conclude $p/\hat{q} \in u$. It then follows that $\sigma \in \Phi$. This completes the proof. □