Appears in 8th European Conference on Logic in Artificial Intelligence (JELIA '2002).

Interpolation Theorems for Nonmonotonic Reasoning Systems

Eyal Amir

Computer Science Division
University of California at Berkeley
Berkeley, CA 94720-1776, USA
eyal@cs.berkeley.edu

Abstract. *Craig's interpolation theorem* [3] is an important theorem known for propositional logic and first-order logic. It says that if a logical formula β logically follows from a formula α, then there is a formula γ, including only symbols that appear in both α, β, such that β logically follows from γ and γ logically follows from α. Such theorems are important and useful for understanding those logics in which they hold as well as for speeding up reasoning with theories in those logics. In this paper we present interpolation theorems in this spirit for three nonmonotonic systems: circumscription, default logic and logic programs with the stable models semantics (a.k.a. answer set semantics). These results give us better understanding of those logics, especially in contrast to their nonmonotonic characteristics. They suggest that some *monotonicity* principle holds despite the failure of classic monotonicity for these logics. Also, they sometimes allow us to use methods for the decomposition of reasoning for these systems, possibly increasing their applicability and tractability. Finally, they allow us to build structured representations that use those logics.

1 Introduction

Craig's interpolation theorem [3] is an important theorem known for propositional logic and first-order logic (FOL). It says that if α, β are two logical formulae and $\alpha \vdash \beta$, then there is a formula $\gamma \in \mathcal{L}(\alpha) \cap \mathcal{L}(\beta)$ such that $\alpha \vdash \gamma$ and $\gamma \vdash \beta$ ("$\vdash$" is the classical logical deduction relation; $\mathcal{L}(\alpha)$ is the language of α (the set of formulae built with the nonlogical symbols of α, $L(\alpha)$)). Such interpolation theorems allow us to break inference into pieces associated with sublanguages of the language of that theory [2, 21], for those formal systems in which they hold. In AI, these properties have been used to speed up inference for constraint satisfaction systems (CSPs), propositional logic and FOL (e.g., [6, 4, 21, 7, 2, 5]) and to build structured representations [4, 1]

In this paper we present interpolation theorems for three nonmonotonic systems: *circumscription* [19], *default logic* [23] and *logic programs* with the Answer Set semantics [12, 10]. In the nonmonotonic setup there are several interpolation theorems for each system, with different conditions for applicability and different form of interpolation. This stands in contrast to classical logic, where Craig's interpolation theorem always holds. Our theorems allow us to use methods for the decomposition of reasoning (a-la [2, 21]) under some circumstances for these systems, possibly increasing their applicability and tractability for structured theories. We list the main theorems that we show in this paper below, omitting some of their conditions for simplicity.

For circumscription we show that, under some conditions, $Circ[\alpha; P; Q] \models \beta$ iff there is some set of formulae $\gamma \subseteq \mathcal{L}(\alpha) \cap \mathcal{L}(\beta)$ such that $\alpha \models \gamma$ and $Circ[\gamma; P; Q] \models \beta$. For example, to answer $Circ[BlockW; block; L(BlockW)] \models on(A, B)$, we can compute this formula $\gamma \in \mathcal{L}(\{block, on, A, B\})$ from $BlockW$ *without applying circumscription*, and then solve $Circ[\gamma; block; L(BlockW)] \models on(A, B)$ (where γ may be significantly smaller than $BlockW$).

For default logic, letting $\alpha \mid\!\sim_D \beta$ mean that every extension of $\langle \alpha, D \rangle$ entails β (*cautious entailment*), we show that, under some conditions, if $\alpha \mid\!\sim_D \beta$, then there is a formula $\gamma \in \mathcal{L}(\alpha \cup D) \cap \mathcal{L}(\beta)$ such that $\alpha \mid\!\sim_D \gamma$ and $\gamma \mid\!\sim_D \beta$. For logic programs we show that if P_1, P_2 are two logic programs and $\varphi \in \mathcal{L}(P_2)$ such that $P_1 \cup P_2 \mid\!\sim^b \varphi$, then there is $\gamma \in \mathcal{L}(P_1) \cap \mathcal{L}(P_2)$ such that $P_1 \mid\!\sim^b \gamma$ and $P_2 \cup \gamma \mid\!\sim^b \varphi$ (here $\mid\!\sim^b$ is the *brave* entailment for logic programs).

This paper focuses on the form of the interpolation theorems that hold for those nonmonotonic logics. We do not address the possible application of these results to the problem of automated reasoning with those logics. Nonetheless, we mention that direct application of those results is possible along the lines already explored for propositional logic and FOL in [2, 21].

No interpolation theorems were shown for nonmonotonic reasoning systems before this paper. Nonetheless, some of our theorems for default logic and logic programs are close to the *splitting theorems* of [17, 26], which have already been used to decompose reasoning for those logics. The main difference between our theorems and those splitting theorems is that the latter change some of the defaults/rules involved to provide the corresponding entailment. Also, they do not talk about an interpolant γ, but rather discuss combining extensions.

Since its debut, the nonmonotonic reasoning line of work has expanded and several textbooks now exist that give a fair view of nonmonotonic reasoning and its uses (e.g., [8]). The reader is referred to those books for background and further details. *[Most proofs are shortened or omitted for lack of space. They are available from the author.]*

2 Logical Preliminaries

In this paper, we use the notion of *logical theory* for every set of axioms in FOL or propositional logic, regardless of whether the set of axioms is deductively closed or not. We use $L(\mathcal{A})$ to denote the signature of $\mathcal{A}$, i.e., the set of non-logical symbols. $\mathcal{L}(\mathcal{A})$ denotes the language of $\mathcal{A}$, i.e., the set of formulae built with $L(\mathcal{A})$. $Cn(\mathcal{A})$ is the set of logical consequences of $\mathcal{A}$ (i.e., those formulae that are valid consequences of $\mathcal{A}$ in FOL). For a first-order structure, M, in L, we write $U(M)$ for the universe of elements of M. For every symbol, s, in L, we write s^M for the interpretation of s in M.

Theorem 1 ([3]). *Let α, β be sentences such that $\alpha \vdash \beta$. Then there is a formula γ involving only nonlogical symbols common to both α and β, such that $\alpha \vdash \gamma$ and $\gamma \vdash \beta$.*

3 Circumscription

3.1 McCarthy's Circumscription: Overview

McCarthy's circumscription [19, 20] is a nonmonotonic reasoning system in which inference from a set of axioms, A, is performed by minimizing the extent of some predicate symbols $\overrightarrow{P}$, while allowing some other nonlogical symbols, $\overrightarrow{Z}$ to vary.

Formally, McCarthy's circumscription formula

$$Circ[A(P,Z);P;Z] = A(P,Z) \wedge \forall p,z\ (A(p,z) \Rightarrow \neg(p < P)) \tag{1}$$

says that in the theory A, with parameter relations and function vectors (sequence of symbols) P, Z, P is a minimal element such that $A(P,Z)$ is still consistent, when we are allowed to vary Z in order to allow P to become smaller.

Take for example the following simple theory: $T = block(B_1) \wedge block(B_2)$. Then, the circumscription of $block$ in T, varying nothing, is

$$Circ[T;block;] = T \wedge \forall p\ [T_{[block/p]} \Rightarrow \neg(p < block)].$$

Roughly, this means that $block$ is a minimal predicate satisfying T. Computing circumscription is discussed in length in [16] and others, and we do not expand on it here. Using known techniques we can conclude

$$Circ[T;block;] \equiv \forall x\ (block(x) \Leftrightarrow (x = B_1 \vee x = B_2))$$

This means that there are no other blocks in the world other than those mentioned in the original theory T.

We give the preferential semantics for circumscription that was given by [15, 20, 9] in the following definition.

Definition 1 ([15]). *For any two models M and N of a theory T we write $M \leq_{P,Z} N$ if the models M, N differ only in how they interpret predicates from P and Z and if the extension of every predicate from P in M is a subset of its extension in N. We write $M <_{P,Z} N$ if for at least one predicate in P the extension in M is a strict subset of its extension in N.*

A model M of T is $\leq_{P,Z}$*-minimal* if there is no model N of T such that $N <_{P,Z} M$.

Theorem 2 ([15]: Circumscript. Semantics). *Let T be a finite set of sentences. A structure M is a model of $Circ[T;P;Z]$ iff M is a $\leq_{P,Z}$-minimal model of T.*

This theorem allows us to extend the definition of circumscription to set of infinite number of sentences. In those cases, $Circ[T;P;Z]$ is defined as the set of sentences that hold in all the $\leq_{P,Z}$-minimal models of T. Theorem 2 implies that this extended definition is equivalent to the syntactic characterization of the original definition (equation (1)) if T is a finite set of sentences. In the rest of this paper, we refer to this extended definition of circumscription, if T is an infinite set of FOL sentences (we will note those cases when we encounter them).

Circumscription satisfied Left Logical Equivalence (LLE): $T \equiv T'$ implies that $Circ[T;P;Z] \equiv Circ[T';P;Z]$. It also satisfies Right Weakening: $Circ[T;P;Z] \models \varphi$ and $\varphi \Rightarrow \psi$ implies that $Circ[T;P;Z] \models \psi$).

3.2 Model Theory

Definition 2. *Let M, N be L-structures, for FOL signature L and language $\mathcal{L}$. We say that N is an* elementary extension *of M (or M is an* elementary substructure *of N), written $M \preceq N$, if $U(M) \subseteq U(N)$ and for every $\varphi(\overrightarrow{x}) \in \mathcal{L}$ and vector of elements $\overrightarrow{a}$ of $U(M)$, $M \models \varphi(\overrightarrow{a})$ iff $N \models \varphi(\overrightarrow{a})$.*

For FOL signatures $L \subseteq L^+$, and for N an L^+-structure, we say that $N \upharpoonright L$ is the *reduct* of N to L, the L-structure with the same universe of elements as N, and the same interpretation as N for those symbols from L^+ that are in L (there is no interpretation for symbols not in L). For A theory T in a language of L^+, let $Cn^L(T)$ be the set of all consequences of T in the language of L.

3.3 Interpolation in Circumscription

In this section we present two interpolation theorems for circumscription. Those theorems hold for both FOL and propositional logic. Roughly speaking, the first (Theorem 4) says that if α nonmonotonically entails β (here this means $Circ[\alpha; P; Q] \models \beta$), then there is $\gamma \subseteq \mathcal{L}(\alpha) \cap \mathcal{L}(\beta \cup P)$ such that α classically entails γ ($\alpha \models \gamma$) and γ nonmonotonically entails β ($Circ[\gamma; P; Q] \models \beta$). In the FOL case this γ can be an infinite set of sentences, and we use the extended definition of Circumscription for infinite sets of axioms for this statement.

The second theorem (Theorem 7) is similar to the first, with two main differences. First, it requires that $L(\alpha) \subseteq (P \cup Q)$. Second, it guarantees that γ as above (and some other restrictions) exists iff α nonmonotonically entails β. This is in contrast to the first theorem that guarantees only that *if* part. The actual technical details are more fine than those rough statements, so the reader should refer to the actual theorem statements.

In addition to these two theorems, we present another theorem that addresses the case of reasoning from the union of theories (Theorem 6). Before we state and prove those theorems, we prove several useful lemmas.

Our first lemma says that if we are given two theories T_1, T_2, and we know the set of sentences that follow from T_2 in the intersection of their languages, then every model of this set of sentences together with T_1 can be extended to a model of $T_1 \cup T_2$.

Lemma 1. *Let T_1, T_2 be two theories, with signatures in L_1, L_2, respectively. Let γ be a set of sentences logically equivalent to $Cn^{L_1 \cap L_2}(T_2)$. For every L_1-structure, $\mathcal{M}$, that satisfies $T_1 \cup \gamma$ there is a $(L_1 \cup L_2)$-structure, $\widehat{\mathcal{M}}$, that is a model of $T_1 \cup T_2$ such that $\mathcal{M} \preceq \widehat{\mathcal{M}} \upharpoonright L_1$.*

Our second lemma says that every $<_{P,Q}$-minimal model of T that is also a model of T' is a $<_{P,Q}$-minimal model of $T \cup T'$.

Lemma 2. *Let T be a theory and P, Q vectors of nonlogical symbols (P includes only predicate symbols). If $\mathcal{M} \models Circ[T; P; Q]$ and $\mathcal{M} \models T \cup T'$, then $\mathcal{M} \models Circ[T \cup T'; P; Q]$.*

The following theorem is central to the rest of our results in this section. It says that when we circumscribe P, Q in $T_1 \cup T_2$ we can replace T_2 by its consequences in $\mathcal{L}(T_1)$, for some purposes and under some assumptions.

Theorem 3. *For T_1, T_2 theories and P, Q vectors of symbols from $L(T_1) \cup L(T_2)$ such that $P \subseteq L(T_1)$, let γ be a set of sentences logically equivalent to $Cn^{L(T_1) \cap L(T_2)}(T_2)$. Then, for all $\varphi \in \mathcal{L}(T_1)$, if $Circ[T_1 \cup T_2; P; Q] \models \varphi$, then $Circ[T_1 \cup \gamma; P; Q] \models \varphi$.*

PROOF We show that for every model of $Circ[T_1 \cup \gamma; P; Q]$ there is a model of $Circ[T_1 \cup T_2; P; Q]$ whose reduct to $L(T_1)$ is an elementary extension of the reduct of the first model to $L(T_1)$.

Let $\mathcal{M}$ be a $L(T_1 \cup T_2)$-structure that is a model of $Circ[T_1 \cup \gamma; P; Q]$. Then, $\mathcal{M} \models T_1 \cup \gamma$. From Lemma 1 we know that there is a $(L_1 \cup L_2)$-structure, $\widehat{\mathcal{M}}$, that is a model of T_2 such that $\mathcal{M} \upharpoonright L(T_1) \preceq \widehat{\mathcal{M}} \upharpoonright L(T_1)$.

Thus, $\widehat{\mathcal{M}}$ is a $\leq_{P,Q}$-minimal model of $T_1 \cup \gamma$. To see this, assume otherwise. Then, there is a model $\mathcal{M}'$ for the signature $L(T_1 \cup T_2)$ such that $\mathcal{M}' <_{P,Q} \widehat{\mathcal{M}}$ and $\mathcal{M}' \models T_1 \cup \gamma$. Take $\mathcal{M}''$ such that the interpretation of all the symbols in $L(T_1)$ is exactly the same as that of $\mathcal{M}'$ and such that the interpretation of all symbols in $L(T_2) \setminus L(T_1)$ is exactly the same as that of $\mathcal{M}$. Then, $\mathcal{M}'' \models T_1 \cup \gamma$ because $T_1 \cup \gamma \subseteq \mathcal{L}(T_1)$. Also, $\mathcal{M}'' <_{P,Q'} \mathcal{M}$, for $Q' = Q \cap L(T_1)$ because $P \subseteq L(T_1)$ and $\mathcal{M}, \widehat{\mathcal{M}}$ agree on the interpretation of symbols in $L(T_1)$ ($\mathcal{M} \upharpoonright L(T_1) \preceq \widehat{\mathcal{M}} \upharpoonright L(T_1)$). Thus, $\mathcal{M}'' <_{P,Q} \mathcal{M}$, since $\mathcal{M}'', \mathcal{M}$ agree on all the interpretation of all symbols in $L(T_2) \setminus L(T_1)$. This contradicts $\mathcal{M} \models Circ[T_1 \cup \gamma; P; Q]$, so $\widehat{\mathcal{M}}$ is a $\leq_{P,Q}$-minimal model of $T_1 \cup \gamma$.

Thus, $\widehat{\mathcal{M}} \models Circ[T_1 \cup \gamma; P; Q]$, and $\widehat{\mathcal{M}} \models T_1 \cup T_2$. From Lemma 2 we get that $\widehat{\mathcal{M}} \models Circ[T_1 \cup T_2; P; Q]$. Now, let $\varphi \in \mathcal{L}(T_1)$ such that $Circ[T_1 \cup T_2; P; Q] \models \varphi$. Then every model of $Circ[T_1 \cup T_2; P; Q]$ satisfies φ. Let $\mathcal{M}$ be a model of $Circ[T_1 \cup \gamma; P; Q]$ in the language $\mathcal{L}(T_1 \cup T_2)$. Then there is $\widehat{\mathcal{M}}$ as above, i.e., $\widehat{\mathcal{M}} \models Circ[T_1 \cup T_2; P; Q]$ and $\mathcal{M} \upharpoonright L(T_1) \preceq \widehat{\mathcal{M}} \upharpoonright L(T_1)$. Thus, $\widehat{\mathcal{M}} \models \varphi$. Since $\mathcal{M} \upharpoonright L(T_1) \preceq \widehat{\mathcal{M}} \upharpoonright L(T_1)$, $\mathcal{M} \models \varphi$. Thus every model of $Circ[T_1 \cup \gamma; P; Q]$ is a model of φ. ■

Theorem 4 (Interpolation for Circumscription 1). *Let T be a theory, P, Q vectors of symbols, and φ a formula. If $Circ[T; P; Q] \models \varphi$, then there is $\gamma \subseteq \mathcal{L}(T) \cap \mathcal{L}(\varphi \cup P)$ such that*

$$T \models \gamma \quad \textit{and} \quad Circ[\gamma; P; Q] \models \varphi.$$

Furthermore, this holds for every γ that is logically equivalent to the consequences of T in $L(T) \cap L(\varphi \cup P)$.

PROOF We use Theorem 3 to find this γ. For T, φ as in the statement of the theorem we define T_1, T_2 as follows. We choose T_1 such that $\varphi \in \mathcal{L}(T_1)$ and $P \subseteq L(T_1)$: Let $T_1 = \{\varphi \vee \neg\varphi\} \cup \tau_1$ for τ_1 a set of tautologies such that $L(\tau_1) = P$. We choose T_2 such that it includes T and has a rich enough vocabulary so that $P, Q \subseteq L(T_1) \cup L(T_2)$. Let $T_2 = T \cup \tau_2$, for τ_2 a set of tautologies such that $L(\tau_2) = Q \setminus L(T_1)$. Let $L_1 = L(T_1)$, $L_2 = L(T_2)$.

Theorem 3 guarantees that if $P \subseteq L_1$ then γ from that theorem satisfies $Circ[T_1 \cup T_2; P; Q] \models \psi \Rightarrow Circ[T_1 \cup \gamma; P; Q] \models \psi$ for every $\psi \in \mathcal{L}(T_1)$. This implies that for every $\psi \in \mathcal{L}(\{\varphi\} \cup \tau_1)$, $Circ[T; P; Q] \models \psi \Rightarrow Circ[\gamma; P; Q] \models \psi$. In particular, $Circ[\gamma; P; Q] \models \varphi$, and this γ satisfies our current theorem. ■

Example 1. For example, if $\alpha = block(A) \wedge block(B) \wedge \forall b\ (clear(b) \Leftrightarrow (block(b) \wedge \forall x \neg on(x, b)))$, $\beta = clear(A)$, $P = on$, $Q = clear$, then one possible interpolant is $\gamma = clear(A) \Leftrightarrow \forall x \neg on(x, A)$ because $\alpha \models \gamma$ and $Circ[\gamma; P; Q] \models \beta$ (because $Circ[clear(A) \Leftrightarrow \forall x \neg on(x, A); on; clear] \equiv \forall x, b \neg(x, b) \wedge clear(A)$).

This theorem does not hold if we require $\gamma \subseteq \mathcal{L}(T) \cap \mathcal{L}(\varphi)$ instead of $\gamma \subseteq \mathcal{L}(T) \cap \mathcal{L}(\varphi \cup P)$. For example, take $\varphi = Q$, $T = \{\neg P \Rightarrow Q\}$, where P, Q are propositional symbols. $Circ[T; P; Q] \models \varphi$. However, every logical consequence of T in $L(\varphi)$ is a tautology. Thus, if the theorem was correct with our changed requirement, γ would be equivalent to $\emptyset$ and $Circ[\gamma; P; Q] \not\models \varphi$.

Theorem 5. *Let T_1, T_2 be two theories, P, Q two vectors of symbols from $L(T_1) \cup L(T_2)$ such that $P \subseteq L(T_1)$ and $P \cup Q \supseteq L(T_2)$. Let γ be a set of sentences logically equivalent to $Cn^{L(T_1) \cap L(T_2)}(T_2)$. Then, for all $\varphi \in \mathcal{L}(T_1)$, if $Circ[T_1 \cup \gamma; P; Q] \models \varphi$, then $Circ[T_1 \cup T_2; P; Q] \models \varphi$.*

From Theorem 3 and Theorem 5 we get the following theorem.

Theorem 6 (Interpolation Between Theories). *Let T_1, T_2 be two theories, P, Q vectors of symbols in $L(T_1) \cup L(T_2)$ such that $P \subseteq L(T_1)$ and $P \cup Q \supset L(T_2)$. Let γ be a set of sentences logically equivalent to $Cn^{L(T_1) \cap L(T_2)}(T_2)$. Then, for every $\varphi \in \mathcal{L}(T_1)$,*

$$Circ[T_1 \cup \gamma; P; Q] \models \varphi \iff Circ[T_1 \cup T_2; P; Q] \models \varphi$$

Theorem 7 (Interpolation for Circumscription 2). *Let T be a theory, P, Q vectors of symbols such that $(P \cup Q) \supseteq L(T)$. Let L_2 be a set of nonlogical symbols. Then, there is $\gamma \in \mathcal{L}(T) \cap \mathcal{L}(L_2 \cup P)$ such that $T \models \gamma$ and for all $\varphi \in \mathcal{L}(L_2)$,*

$$Circ[T; P; Q] \models \varphi \iff Circ[\gamma; P; Q] \models \varphi.$$

Furthermore, this γ can be logically equivalent to the consequences of T in $L(T) \cap (L_2 \cup P)$.

PROOF Let T_1 be a set of tautologies such that $L(T_1) = L_2 \cup P$. Also, let $T_2 = T \cup \tau_2$, for τ_2 a set of tautologies such that $L(\tau_2) = Q \setminus L(T_1)$. Let $L_1 = L(T_1)$, $L_2 = L(T_2)$. Theorem 6 guarantees that γ from that theorem satisfies $Circ[\gamma; P; Q] \models \psi \iff Circ[T; P; Q] \models \psi$ for every $\psi \in \mathcal{L}_1 = L_2 \cup P$. ■

Example 2. If α, β and P are taken as in Example 1, and $Q = \{clear, block, A, B\}$, then one interpolant for Theorem 7 is $\gamma = clear(A) \Leftrightarrow \forall x \neg on(x, A)$.

The theorems we presented are for parallel circumscription, where we minimize all the minimized predicates in parallel without priorities. The case of prioritized circumscription is outside the scope of this paper.

4 Default Logic

In this section we present interpolation theorems for *propositional* default logic. We also assume that the signature of our propositional default theories is finite (this also implies that our theories are finite).

4.1 Reiter's Default Logic: Overview

In Reiter's default logic [23] one has a set of facts W (in either propositional or FOL) and a set of defaults D (in a corresponding language). Defaults in D are of the form $\frac{\alpha:\beta_1,\dots,\beta_n}{\delta}$ with the intuition that if α is proved, and $\beta_1,...,\beta_n$ are consistent (throughout the proof), then δ is proved. α is called the *prerequisite*, $pre(d) = \{\alpha\}$; $\beta_1,...,\beta_n$ are the *justifications*, $just(d) = \{\beta_1,...,\beta_n\}$ and δ is the *consequent*, $cons(d) = \{\delta\}$. We use similar notation for sets of defaults (e.g., $cons(D) = \bigcup_{d\in D} cons(d)$).

Definition 3. *An* extension *of $\langle W, D\rangle$ is a set of sentences E that satisfies W, follows the defaults in D, and is minimal. More formally, E is an extension if it is minimal (as a set) such that $\Gamma(E) = E$, where we define $\Gamma(S_0)$ to be S, a minimal set of sentences such that*

1. *$W \subseteq S$; $S = Cn(S)$.*
2. *For all $\frac{\alpha:\beta_1,\dots,\beta_n}{\delta} \in D$ if $\alpha \in S$ and $\forall i\ \neg\beta_i \notin S_0$, then $\delta \in S$.*

The following theorem provides an equivalent definition that was shown in [18, 24] and others. A set of defaults, $\mathbb{D}$ is grounded in a set of formulae W iff for all $d \in \mathbb{D}$, $pre(d) \in Cn_{Mon(\mathbb{D})}(W)$, where $Mon(\mathbb{D}) = \{\frac{pre(d)}{cons(d)} \mid d \in \mathbb{D}\}$.

Theorem 8 (Extensions in Terms of Generating Defaults). *A set of formulae E is an extension of a default theory $\langle W, D\rangle$ iff $E = Cn(W \cup \{cons(d) \mid d \in D'\})$ for a minimal set of defaults $D' \subseteq D$ such that*

1. *D' is grounded in W and*
2. *for every $d \in D, d \in D'$ if and only if $pre(d) \in Cn(W \cup cons(D'))$ and every $\psi \in just(d)$ satisfies $\neg\psi \notin Cn(W \cup cons(D'))$.*

Every minimal set of defaults $D' \subseteq D$ as mentioned in this theorem is said to be a set of *generating defaults*.

Normal defaults are defaults of the form $\frac{\alpha:\beta}{\beta}$. These defaults are interesting because they are fairly intuitive in nature (if we proved α then β is proved unless previously proved inconsistent). A default theory is *normal*, if all of its defaults are normal.

We define $W \mid\!\sim_D \varphi$ as cautious entailment sanctioned by the defaults in D, i.e., φ follows from every extension of $\langle W, D\rangle$. We define $W \mid\!\sim_D^b \varphi$ as brave entailment sanctioned by the defaults in D, i.e., φ follows from at least one extension of $\langle W, D\rangle$.

4.2 Interpolation in Default Logic

In this section we present several flavors of interpolation theorems, most of which are stated for cautious entailment.

Theorem 9 (Interpolation for Cautious DL 1). *Let $T = \langle W, D\rangle$ be a propositional default theory and φ a propositional formula. If $W \mid\!\sim_D \varphi$, then there are γ_1, γ_2 such that $\gamma_1 \in \mathcal{L}(W) \cap \mathcal{L}(D \cup \{\varphi\})$, $\gamma_2 \in \mathcal{L}(W \cup D) \cap \mathcal{L}(\varphi)$ and all the following hold:*

$$W \models \gamma_1 \quad \gamma_1 \mid\!\sim_D \gamma_2 \quad \gamma_2 \models \varphi \quad W \mid\!\sim_D \gamma_2 \quad \gamma_1 \mid\!\sim_D \varphi$$

PROOF Let γ_1 be the set of consequences of W in $\mathcal{L}(D \cup \{\varphi\}) \cap \mathcal{L}(W)$. Let $\mathbb{E}$ be the set of extensions of $\langle W, D\rangle$ and $\mathbb{E}'$ the set of extensions of $\langle \gamma_1, D\rangle$. We show that every extension $E' \in \mathbb{E}'$ has an extension $E \in \mathbb{E}$ such that $Cn(E' \cup W) = Cn(E)$. This will show that γ_1 is as needed.

Take $E' \in \mathbb{E}'$ and define $E_0 = Cn(E' \cup W)$. We assume that $L(E') \subseteq L(\mathbb{D})$ because otherwise we can take a logically equivalent extension whose sentences are in $\mathcal{L}(\mathbb{D})$. We show that E_0 satisfies the conditions for extensions of $\langle W, D\rangle$ (Definition 3). The first condition holds by definition of E_0. The second condition holds because every default that is consistent with E_0 is also consistent with E' and vice versa.

For the first direction (every default that is consistent with E_0 is also consistent with E'), let $\frac{\alpha:\beta_1,\ldots,\beta_n}{\delta} \in D$ be such that $\alpha \in E_0$. We show that $\alpha \in E'$. By definition, $\alpha \in \mathcal{L}(D)$. $\alpha \in E_0$ implies that $E' \cup W \models \alpha$ because $Cn(E' \cup W) = E_0$. Using the deduction theorem for propositional logic we get $W \models E' \Rightarrow \alpha$ (taking E' here to be a finite set of sentences that is logically equivalent to E' in $\mathcal{L}(\mathbb{D})$ (there is such a finite set because we assume that $L(\mathbb{D})$ is finite)). Using Craig's interpolation theorem for propositional logic, there is $\gamma \in \mathcal{L}(W) \cap \mathcal{L}(E' \Rightarrow \alpha)$ such that $W \models \gamma$ and $\gamma \models E' \Rightarrow \alpha$. However, this means that $\gamma_1 \models \gamma$, by the way we chose γ_1. Thus $\gamma_1 \models E' \Rightarrow \alpha$. Since $E' \subseteq \gamma_1$ we get that $E' \models \alpha$. Since $E' = Cn(E')$ we get that $\alpha \in E'$. The case is similar for δ: if $\delta \in E_0$ then $\delta \in E'$ by the same argument as given above for $\alpha \in E_0 \Rightarrow \alpha \in E'$. Finally, if $\forall i\ \neg\beta_i \notin E_0$ then $\forall i\ \neg\beta_i \notin E'$ because $E' \subseteq E_0$. The opposite direction (every default that is consistent with E' is also consistent with E_0) is similar to the first one.

Thus, E_0 satisfies those two conditions. However, it is possible that E_0 is not a minimal such set of formulae. If so, Theorem 8 implies that there is a strict subset of the generating defaults of E_0 that generate a different extension. However, we can apply this new set of defaults to generate an extension that is smaller than E', contradicting the fact that E' is an extension of $\langle \gamma_1, D\rangle$.

Now, if φ logically follows in all the extensions of $\langle W, D\rangle$ then it must also follow from every extension of $\langle \gamma_1, D\rangle$ together with W. Let $\Lambda = E_1 \vee \ldots \vee E_n$, for $E_1, \ldots, E_n$ the (finite) set of (logically non-equivalent) extensions of $\langle W, D\rangle$ (we have a finite set of those because $L(W) \cup L(D)$ is finite). Then, $\Lambda \models \varphi$. Take $\gamma_2 \in \mathcal{L}(\Lambda) \cap \mathcal{L}(\varphi)$ such that $\Lambda \models \gamma_2$ and $\gamma_2 \models \varphi$, as guaranteed by Craig's interpolation theorem (Theorem 1). These γ_1, γ_2 are those promised by the current theorem: $W \models \gamma_1$, $\gamma_2 \models \varphi$, $W \mid\!\sim_D \gamma_2$, $\gamma_1 \mid\!\sim_D \gamma_2$ and $\gamma_1 \mid\!\sim_D \varphi$. ■

Theorem 10 (Interpolation for Cautious DL 2). *Let $T = \langle W, D\rangle$ be a propositional default theory and φ a propositional formula. If $W \mid\!\sim_D \varphi$, then there are $\gamma_1, \gamma_2 \in \mathcal{L}(W) \cap \mathcal{L}(D)$, and all the following hold:*

$$W \models \gamma_1 \quad \gamma_1 \mid\!\sim_D \gamma_2 \quad \{\gamma_2\} \cup W \models \varphi \quad W \mid\!\sim_D \gamma_2$$

Example 3. If $\alpha = wet \wedge (wet \Rightarrow mud)$, $\beta = (rain \Rightarrow clouds) \Rightarrow clouds$, $D = \{\frac{mud:rain}{rain}\}$, then one possible pair of interpolants is $\gamma_1 = mud$, $\gamma_2 = rain$.

For another example, if the only objects are A, B, and $\alpha = \mathit{block}(A) \wedge \mathit{block}(B) \wedge \forall b\ (\mathit{clear}(b) \Leftrightarrow (\mathit{block}(b) \wedge \forall x \neg \mathit{on}(x, b)))$, $\beta = \mathit{clear}(A)$ and $D = \{\frac{:\neg on(a,b)}{\neg on(a,b)}\}$, then one possible pair of interpolants is $\gamma_1 = \mathit{TRUE}$, $\gamma_2 = \neg \mathit{on}(B, A) \wedge \neg \mathit{on}(A, A)$. To see that γ_1, γ_2 satisfy Theorem 10 notice that $\langle \alpha, D \rangle$ has one extension: $E = \{block(A), block(B), \neg on(A, B), \neg on(B, A), \neg on(A, A), \neg on(B, B), clear(A), clear(B)\}$.

Corollary 1. *Let $\langle W, D \rangle$ be a default theory and φ a formula. If $W \mathrel{|\!\!\sim}_D \varphi$, then there is a set of formulae, $\gamma \in \mathcal{L}(W \cup D) \cap \mathcal{L}(\varphi)$ such that $W \mathrel{|\!\!\sim}_D \gamma$ and $\gamma \mathrel{|\!\!\sim}_D \varphi$.*

We do not get stronger interpolation theorems for prerequisite-free normal default theories. [14] provided a modular translation of normal default theories with no prerequisites into circumscription, but Theorem 4 does not lead to better results. In particular, the counter example that we presented after Theorem 4 can be massaged to apply here too.

Theorem 11 (Interpolation Between Default Extensions). *Let $\langle W_1, D_1 \rangle, \langle W_2, D_2 \rangle$ be default theories such that $L(cons(D_2)) \cap L(pre(D_1) \cup just(D_1) \cup W_1) = \emptyset$. Let φ be a formula such that $\varphi \in \mathcal{L}(W_2 \cup D_2)$. If there is an extension E of $\langle W_1 \cup W_2, D_1 \cup D_2 \rangle$ in which φ holds, then there is a formula $\gamma \in \mathcal{L}(W_1 \cup D_1) \cap \mathcal{L}(W_2 \cup D_2)$, an extension E_1 of $\langle W_1, D_1 \rangle$ such that $Cn(E_1) \cap \mathcal{L}(W_2 \cup D_2) = \gamma$, and an extension E_2 of $\langle W_2 \cup \{\gamma\}, D_2 \rangle$ such that $E_2 \models \varphi$.*

It is interesting to notice that the reverse direction of this theorem does not hold. For example, if we have two extensions E_1, E_2 as in the theorem statement, it is possible that E_1 uses a default with justification β, but $W_2 \models \neg\beta$. Strengthening the condition of the theorem, i.e., demanding that $L(W_2 \cup cons(D_2)) \cap L(pre(D_1) \cup just(D_1) \cup W_1) = \emptyset$, is not sufficient either. For example, if D_1 includes two defaults $d_1 = \frac{:}{a \Rightarrow \neg\beta}$, and $d_2 = \frac{:\beta}{\varphi}$, $W_1 = \emptyset$, D_2 includes no defaults and $W_2 = \{a\}$ then there is no extension of $\langle W_1 \cup W_2, D_1 \cup D_2 \rangle$ that implies φ, for $\varphi = \{c\}$.

Further strengthening the conditions of the theorem gives the following:

Theorem 12 (Reverse Direction of Theorem 11). *Let $\langle W_1, D_1 \rangle, \langle W_2, D_2 \rangle$ be default theories such that $L(W_2 \cup cons(D_2)) \cap L(D_1 \cup W_1) = \emptyset$. Let φ be a formula such that $\varphi \in \mathcal{L}(W_2 \cup D_2)$. There is an extension E of $\langle W_1 \cup W_2, D_1 \cup D_2 \rangle$ in which φ holds only if there is a formula $\gamma \in \mathcal{L}(W_1 \cup D_1) \cap \mathcal{L}(W_2 \cup D_2)$, an extension E_1 of $\langle W_1, D_1 \rangle$ such that $Cn(E_1) \cap \mathcal{L}(W_2 \cup D_2) = \gamma$, and an extension E_2 of $\langle W_2 \cup \{\gamma\}, D_2 \rangle$ such that $E_2 \models \varphi$.*

Corollary 2 (Interpolation for Brave DL). *Let $\langle W_1, D_1 \rangle, \langle W_2, D_2 \rangle$ be default theories such that $L(cons(D_2)) \cap L(pre(D_1) \cup just(D_1) \cup W_1) = \emptyset$. Let φ be a formula such that $\varphi \in \mathcal{L}(W_2 \cup D_2)$. If $W_1 \cup W_2 \mathrel{|\!\!\sim}^b_{D_1 \cup D_2} \varphi$, then there is a formula, $\gamma \in \mathcal{L}(W_1 \cup D_1) \cap \mathcal{L}(W_2 \cup D_2)$, such that $W_1 \mathrel{|\!\!\sim}^b_{D_1} \gamma$ and $W_2 \cup \{\gamma\} \mathrel{|\!\!\sim}^b_{D_2} \varphi$.*

Finally, Corollary 2 and Theorem 11 are similar to the *splitting theorem* of [26], which is provided for default theories with $W = \emptyset$ (there is a modular translation that converts every default theory to one with $W = \emptyset$). A *splitting set* for a set of defaults D is a subset A of $L(D)$ such that $pre(D), just(D), cons(D) \subseteq \mathcal{L}(A) \cup \mathcal{L}(L(D) \setminus A)$ and $\forall d \in D\ (cons(d) \notin \mathcal{L}(L(D) \setminus A) \Rightarrow L(d) \subseteq A)$. Let $B = L(D) \setminus A$. The base of D relative to A is $b_A(D) = \{d \in D \mid L(d) \subseteq A)$. For a set of sentences $X \subseteq \mathcal{L}(A)$, we define $e_A(D, X)$ to be

$$\left\{ \frac{\bigwedge(\{a_i\}_{i\leq n} \cap \mathcal{L}(B)) : \{b_i\}_{i\leq m} \cap \mathcal{L}(B)}{c} \;\middle|\; \begin{array}{l} \frac{\bigwedge_{i=1}^{n} a_i : b_1, \ldots, b_m}{c} \in D \setminus b_A(D), \\ \forall i \leq n(a_i \in \mathcal{L}(A) \Rightarrow a_i \in Cn^A(X)), \\ \forall i \leq m(\neg b_i \notin Cn^A(X)) \end{array} \right\}$$

Theorem 13 ([26]). *Let A be a splitting set for a default theory D over $\mathcal{L}(U)$. A set E of formulae is a consistent extension of D iff $E = Cn^{L(D)}(X \cup Y)$, for some consistent extension X of $b_a(D)$ over $\mathcal{L}(A)$ and Y a consistent extension of $e_A(D, X)$ over $\mathcal{L}(L(D) \setminus A)$.*

Roughly speaking, this theorem finds an extension X of the *base* ($b_A(D)$) and converts $D \setminus b_A(D)$ using this X into a theory $e_A(D, X)$. Then, an extension Y for $e_A(D, X)$ completes the extension for D if $X \cup Y$ is consistent. In contrast, our theorem does not change $D \setminus b_A(D)$, but it is somewhat weaker, in that it only provides a necessary condition for $D \mid\!\sim^b \varphi$. (however, notice that this *weaker* form is typical for interpolation theorems).

5 Logic Programs

In this section we provide interpolation theorems for logic programs with the stable models semantics. We use the fact the logic programs are a special case of default logic, and the results are straightforward. An *extended disjunctive logic program* [10–12, 22] is a set of *rules*. Each rule, r, is written as an expression of the form

$$L_1|...|L_l \leftarrow A_1, ..., A_n, notB_1, ..., notB_m$$

where $L_1, ..., L_l, A_1, ..., A_n, B_1, ..., B_m$ are literals, that is, atomic formulae or their (classic) negations, $L_1, ..., L_l$ are the *head literals*, $head(r)$, $A_1, ..., A_n$ are the *positive subgoals*, $pos(r)$, and $B_1, ..., B_m$ are the *negated subgoals*, $neg(r)$.

[25] showed that disjunctive logic programs (no classic negation) with the stable model semantics can be translated to prerequisite-free default theories as follows:

1. For a rule $A_1|...|A_l \leftarrow B_1, ..., B_m,\ not\ C_1, ...,\ not\ C_n$ in P, we get the default
$$\frac{: \neg C_1, ..., \neg C_n}{B_1 \wedge ... \wedge B_m \Rightarrow A_1 \vee ... \vee A_l}$$
2. For each atom A appearing in P, we get the default $\frac{:\neg A}{\neg A}$

Each stable model of P is the set of atoms in some extension of D_P, and the set of atoms in an extension of D_P is a stable model of P (notice that, in general, an extension of D_P can include sentences that are not atoms and are not subsumed by atoms in that extension). [25] provide a similar translation to extended disjunctive logic programs by first translating those into disjunctive logic programs (a literal $\neg A$ is translated to a new symbol, A'), showing that a similar property holds for this class of programs.

We define $P \mid\!\sim \varphi$ as cautious entailment sanctioned from the logic program P, i.e., φ follows from stable model of P. We define $P \mid\!\sim^b \varphi$ as brave entailment sanctioned from the logic program P, i.e., φ follows from at least one stable model of P.

From the translation above we get the following interpolation theorems.

Theorem 14 (Interpolation for Stable Models (Cautious)). *Let P be a logic program and let φ be a formula such that $P \mid\!\sim \varphi$. Then, there is a formula $\gamma \in \mathcal{L}(P) \cap \mathcal{L}(\varphi)$ such that $P \mid\!\sim \gamma$ and $\gamma \models \varphi$.*

PROOF Follows if γ_2 in Theorem 9 corresponding to our needed γ. ■

Theorem 15 (Interpolation for Stable Models (Brave)). *Let P_1, P_2 be logic programs such that $head(P_2) \cap body(P_1) = \emptyset$. Let $\varphi \in \mathcal{L}(P_2)$ be a formula such that $P_1 \cup P_2 \mid\!\sim^b \varphi$. Then, there is a formula $\gamma \in \mathcal{L}(P_1) \cap \mathcal{L}(P_2)$ such that $P_1 \mid\!\sim^b \gamma$ and $\gamma \cup P_2 \mid\!\sim^b \varphi$.*

PROOF Follows from the reduction to default logic and Corollary 2. ■

The last theorem is similar to the *splitting theorem* of [17]. This theorem finds an answer set X of the *bottom* (P_1) and converts P_2 using this X into a program P_2'. Then, an answer set Y for P_2' completes the answer set for $P_1 \cup P_2$ if $X \cup Y$ is consistent. In contrast, our theorem does not change P_2, but it is somewhat weaker, in that it does only provides a necessary condition for $P_1 \cup P_2 \mid\!\sim^b \varphi$ (this is the typical form of an interpolation theorem).

6 Summary

We presented interpolation theorems that are applicable to the nonmonotonic systems of circumscription, default logic and Answer Set Programming (a.k.a. Stable Models Semantics). These results are somewhat surprising and revealing in that they show particular structure for the nonmonotonic entailments associated with the different systems. They promise to help in reasoning with larger systems that are based on these nonmonotonic systems.

Several questions remain open. First, γ promised by our theorems is not always finite (in the FOL case). This is in contrast to classical FOL, where the interpolant is always of finite length. What conditions guarantee that it is finite in our setup? We conjecture that this will require the partial order involved in the circumscription to be *smooth*. Second, are there better interpolation theorems for the prioritized case of those systems? Also, what is the shape of the interpolation theorems specific for prerequisite-free semi-normal defaults? Further, our results for default logic and logic programs are propositional. How do they extend to the FOL case?

Finally, the theorems for default logic and Logic Programming promise that $\alpha \mid\!\sim_D \beta$ implies the existence of γ such that $\alpha \mid\!\sim_D \gamma$ and $\gamma \mid\!\sim_D \beta$. However, we do not know that the other direction holds, i.e., that the existence of γ such that $\alpha \mid\!\sim_D \gamma$ and $\gamma \mid\!\sim_D \beta$ implies that $\alpha \mid\!\sim_D \beta$. Can we do better than Theorem 12 for different cases?

7 Acknowledgments

I wish to thank Esra Erdem, Vladimir Lifschitz and Leora Morgenstern for reading and commenting on different drafts of this manuscript. This research was supported by DARPA grant N66001-00-C-8018 (RKF program).

References

1. E. Amir. (De)composition of situation calculus theories. In *Proc. AAAI '00*, pages 456–463. AAAI Press/MIT Press, 2000.
2. E. Amir and S. McIlraith. Paritition-based logical reasoning. In *Proc. KR '2000*, pages 389–400, 2000.
3. W. Craig. Linear reasoning. a new form of the Herbrand-Gentzen theorem. *J. of Symbolic Logic*, 22:250–268, 1957.
4. A. Darwiche. Model-based diagnosis using structured system descriptions. *Journal of AI Research*, 8:165–222, 1998.
5. R. Dechter. Bucket elimination: A unifying framework for reasoning. *Artificial Intelligence*, 113(1–2):41–85, 1999.
6. R. Dechter and J. Pearl. Tree Clustering Schemes for Constraint Processing. In *Proc. AAAI '88*, 1988.
7. R. Dechter and I. Rish. Directional resolution: The Davis-Putnam procedure, revisited. In *Proc. KR '94*, pages 134–145. Morgan Kaufmann, 1994.
8. J. D.M. Gabbay, C.J.Hogger, editor. *Handbook of Logic in Artificial Intelligence and Logic Programming, Vol. 3: Nonmonotonic Reasoning and Uncertain Reasoning*. Oxford, 1993.
9. D. Etherington. *Reasoning with incomplete Information*. PhD thesis, University of British Columbia, 1986.
10. M. Gelfond and V. Lifschitz. The stable model semantics for logic programming. In *5th International Conference on Logic Programming*, pages 1070–1080, 1988.
11. M. Gelfond and V. Lifschitz. Logic Program with Classical Negation. In D. H. D. Warren and P. Szeredi, editors, *7th Int. Conf. on Logic Programming*, pages 579–597. MIT, 1990.
12. M. Gelfond and V. Lifschitz. Classical Negation in Logic Programs and Disjunctive Databases. *New Generation Computing*, 9, 1991.
13. W. Hodges. *A shorter model theory*. Cambridge U. Press, 1997.
14. T. Imielinski. Results on translating defaults to circumscription. *Artificial Intelligence*, 32(1):131–146, Apr. 1987.
15. V. Lifschitz. Computing circumscription. In *Proc. of IJCAI-85*, pages 121–127, 1985.
16. V. Lifschitz. Circumscription. In D. Gabbay, C.J.Hogger, and J.A.Robinson, editors, *H.B. of Logic in Artificial Intelligence and Logic Programming, Vol. 3*. Oxford U. Press, 1993.
17. V. Lifschitz and H. Turner. Splitting a logic program. In *Proc. 11th Int'l Conf. on Logic Programming*, pages 23–37. MIT Press, 1994.
18. V. M. Marek and M. Truszczyński. *Nonmonotonic Logics; Context-Dependent Reasoning*. Springer Verlag, Berlin-Heidelberg-New York, 1st edition, 1993.
19. J. McCarthy. Circumscription—A Form of Non-Monotonic Reasoning. *Artificial Intelligence*, 13:27–39, 1980.
20. J. McCarthy. Applications of Circumscription to Formalizing Common Sense Knowledge. *Artificial Intelligence*, 28:89–116, 1986.
21. S. McIlraith and E. Amir. Theorem proving with structured theories. In *IJCAI '01*, pages 624–631, 2001.
22. T. Przymusiński. Stable semantics for disjunctive programs. *New Generation Computing*, 9:401–424, 1991.
23. R. Reiter. A logic for default reasoning. *Artificial Intelligence*, 13 (1–2):81–132, 1980.
24. V. Risch and C. Schwind. Tableau-based characterization and theorem proving for default logic. *Journal of Automated Reasoning*, 13:223–242, 1994.
25. C. Sakama and K. Inoue. Relating disjunctive logic programs to default theories. In *LP-NMR'93*, pages 266–282, 1993.
26. H. Turner. Splitting a default theory. In *Proc. AAAI '96*, pages 645–651, 1996.