Circonscriptions : cas de complétude et d'incomplétude

Philippe Besnard, Jean Houdebine, Raymond Rolland

UNITÉ DE RECHERCHE
INRIA-RENNES

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P.105
78153 Le Chesnay Cedex
France
Tél.:(1) 39 63 55 11

Rapports de Recherche

N° 745

CIRCONSCRIPTIONS : CAS DE COMPLETUDE ET D'INCOMPLETUDE

Philippe BESNARD
Jean HOUDEBINE
Raymond ROLLAND

NOVEMBRE 1987

IRISA

INSTITUT DE RECHERCHE EN INFORMATIQUE
ET SYSTÈMES ALÉATOIRES

Campus Universitaire de Beaulieu
Avenue du Général Leclerc
35042 - RENNES CÉDEX
FRANCE
Tél. : (99) 36.20.00
Télex : UNIRISA 95 0473 F

Circonscriptions: cas de complétude et d'incomplétude

Circumscriptions: completeness and incompleteness cases

Publication Interne 373 – Septembre 1987 – 16 Pages

Philippe Besnard, Jean Houdebine et Raymond Rolland

Septembre 1987

Ph. Besnard: IRISA – J. Houdebine et R. Rolland: Département de Mathématiques

Campus de Beaulieu 35042 Rennes Cédex

Résumé

Le problème de complétude pour différents types de circonscriptions est abordé selon une approche unifiée menant à un théorème très puissant qui permet de retrouver tous les résultats partiels de complétude déjà connus. L'intérêt et la généralité de cette approche et du théorème auquel elle conduit sont illustrés par l'obtention d'un résultat d'incomplétude pour ces circonscriptions, résultat qui contient tous les autres publiés antérieurement.

Abstract

A new approach to the completeness problem for circumscriptions is presented, that leads to a powerful theorem from which all known partial completeness results for circumscriptions can be proved. Also, the theorem yields an incompleteness result for circumscriptions stronger than the ones published elsewhere in the literature.

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE
(L.A. 227)
UNIVERSITÉ DE RENNES 1 I.N.S.A. DE RENNES

INSTITUT NATIONAL DE RECHERCHE
EN INFORMATIQUE ET EN AUTOMATIQUE
(LABORATOIRE DE RENNES)

1 Introduction

Le terme "circonscription" désigne un principe de minimisation logique qui trouve son origine dans les systèmes de raisonnement informatiques [McCarthy 1980] et dont l'objectif est la minimisation de prédicats sous certaines conditions exprimées à l'aide de formules de la logique des prédicats (du premier ordre sauf mention contraire).

L'idée est de minimiser un prédicat relativement à l'ordre représenté par l'inclusion ensembliste (un prédicat est assimilé à un ensemble car un symbole de prédicat – supposé unaire pour simplifier – est interprété dans un modèle par un sous-ensemble du domaine de ce modèle).

En fait, circonscrire le prédicat P dans la théorie T revient à déterminer une extension de T dans laquelle la propriété P (pour simplifier, P est supposé unaire) n'est attribuée qu'aux seuls individus pour lesquels cela est nécessaire. Nécessaire signifie qu'il s'agit de préserver la cohérence de T lorsqu'elle existe.

Par exemple, circonscrire P dans la formule $P(a) \wedge Q(b)$ doit permettre de conclure que P n'est satisfait que par a, donc $\forall x\ P(x) \Rightarrow x = a$.

Le procédé technique employé pour circonscrire un prédicat dans une théorie consiste à compléter la logique du premier ordre par un schéma d'axiomes particulier appelé schéma de circonscription.

Définition 1: Circonscription prédicative

Le schéma pour la circonscription du prédicat P dans la théorie finiment axiomatisable T est de la forme

$$\{\mathrm{T}[\phi_P(\ ,\vec{y})] \wedge \forall \vec{x}\,(\phi_P(\vec{x},\vec{y}) \Rightarrow P(\vec{x}))\} \Rightarrow \{\forall \vec{x}\,(P(\vec{x}) \Rightarrow \phi_P(\vec{x},\vec{y}))\}$$

où
- $\mathrm{T}[\phi_P(\ ,\vec{y})]$ résulte du remplacement (sur la base $P(\vec{x}) \Leftrightarrow \phi_P(\vec{x},\vec{y})$) de P par $\phi_P(\ ,\vec{y})$ partout dans $\mathrm{T}[P]$ (qui est une axiomatisation finie de T)
- $\phi_P(\vec{x},\vec{y})$ peut être toute formule dont les seules variables libres qui se retrouvent liées dans $\mathrm{T}[\phi_P(\ ,\vec{y})]$ appartiennent à $\vec{x}$

Notation: Par la suite, $C_{\mathbf{T}}\{P\} \vdash A$ signifie que la formule A est dérivable dans la logique du premier ordre de $\mathbf{T}$ et du schéma d'axiomes pour la circonscription de P dans $\mathbf{T}$. Par ailleurs, dans les deux premières sections, l'expression circonscription prédicative est abrégée par le seul terme de circonscription.

Exemple 1: Reprenons en la notant $\mathbf{T}_1$ la théorie dont le seul axiome est $P(a) \wedge Q(b)$. Parmi toutes les instances du schéma de circonscription de P dans $\mathbf{T}_1$, considérons celle que donne

$\phi_P(x)$ valant $x = a$

Dans ce cas, en effet, $\mathbf{T}_1[\phi_P]$ devient

$a = a \wedge Q(b)$

qu'il est facile de déduire de $\mathbf{T}_1$. D'autre part, $\forall x\ (\phi_P(x) \Rightarrow P(x))$ vaut

$\forall x\ (x = a \Rightarrow P(x))$

qui se déduit aussi de $\mathbf{T}_1$. Il est maintenant aisé de dériver de cette instance particulière du schéma de circonscription la conclusion désirée

$\forall x\ (P(x) \Rightarrow x = a)$

Cet exemple suscite un commentaire. Tout d'abord, identifier $\phi_P(x)$ à $x = a$, la formule substituée à P est clairement optimal dans la mesure où substituer $x = b$ pour $\phi_P(x)$ à P est d'un nettement moindre intérêt (toutefois cet intérêt n'est pas nul car le résultat correspondant montre que si $a = b$ alors la circonscription de P dans $\mathbf{T}_1$ rend P vrai seulement en a). En fait, $x = a$ pour $\phi_P(x)$ est une définition nouvelle de P car elle a exactement même arité que P et ne contient pas le symbole P lui-même. Cette situation est bien sûr idéale, mais pas générale comme le montrera la dernière section.

L'exemple ci-dessus permet également d'illustrer le cas de formules $\phi_P(\vec{x},\vec{y})$ qui comportent effectivement des variables libres $\vec{y}$ restant libres dans $\mathbf{T}[\phi_P(\ ,\vec{y})]$. Pour cela, il suffit de considérer

$\phi_P(x,y)$ valant $x = y$

Alors $\mathbf{T}_1[\phi_P(\ ,y)]$ devient

$a = y \wedge Q(b)$

D'autre part, $\forall x\ (\phi_P(x,y) \Rightarrow P(x))$ vaut

$\forall x\ (x = y \Rightarrow P(x))$

Est alors obtenue comme instance du schéma de circonscription, la formule

$$\{(a = y \wedge Q(b) \wedge\ \forall x\ (x = y \Rightarrow P(x))\}\ \Rightarrow\ \{\forall x\ (P(x) \Rightarrow x = y)\}$$

Mais dans $\mathbf{T}_1$ cette formule est équivalente à

$$a = y \Rightarrow \{\forall x \ (P(x) \Rightarrow x = y)\}$$

qui est aussi équivalente à la formule $\forall x \ (P(x) \Rightarrow x = a)$ obtenue plus haut.

Comme il l'a été dit en introduction, la circonscription vise à minimiser un prédicat P dans une théorie $\mathbf{T}$ conformément à l'ordre induit par l'inclusion ensembliste. La théorie des modèles développée pour la circonscription suit donc tout simplement cette idée et correspond ainsi, dans le cas de la circonscription d'un prédicat P dans une théorie $\mathbf{T}$, à l'étude des modèles de $\mathbf{T}$ ayant le moins possible d'individus à satisfaire P.

Définition 2: Modèles minimaux

Un modèle $\mathbf{m}$ d'une théorie $\mathbf{T}$ est un modèle minimal de $\mathbf{T}$ relativement à un prédicat P (ou plus brièvement un modèle P-minimal de $\mathbf{T}$) ssi

il n'existe aucun modèle $\mathbf{m}'$ de $\mathbf{T}$ tel que $\mathbf{m}' <_P \mathbf{m}$

où $\mathbf{m}' <_P \mathbf{m}$ ssi $|P|_{\mathbf{m}'} \subset |P|_{\mathbf{m}}$ avec $\mathbf{m}$ et $\mathbf{m}'$ identiques par ailleurs

Notation: $|P|_{\mathbf{m}}$ dénote la réalisation du prédicat P dans le modèle $\mathbf{m}$ c'est-à-dire un sous-ensemble de $|\mathbf{m}|^n$ où $|\mathbf{m}|$ est le domaine de ce modèle.

Il s'ensuit des définitions 1 et 2 que si un modèle $\mathbf{m}$ d'une théorie $\mathbf{T}$ est minimal relativement à un prédicat P alors $\mathbf{m}$ ne peut falsifier aucune instance du schéma de circonscription de P dans $\mathbf{T}$: autrement, il existerait une formule $\phi_P(\vec{x}, \vec{y})$ à partir de laquelle pourrait être définit un modèle $\mathbf{n}$ de $\mathbf{T}$ tel que $\mathbf{n} <_P \mathbf{m}$. C'est ainsi que se démontre le théorème établissant que la circonscription prédicative est consistante relativement à la théorie des modèles minimaux.

Théorème 1: Consistance de la circonscription (McCarthy)

Si $C_{\mathbf{T}}\{P\} \vdash A$ alors A est vrai dans tout modèle de $\mathbf{T}$ minimal relativement à P

Le cas de la réciproque, c'est-à-dire de la complétude de la circonscription relativement à la théorie des modèles minimaux, est tout différent et constitue le sujet des sections suivantes.

2 Les modèles minimaux comme sous-modèles définissables

Le problème de la complétude de la circonscription tel qu'il a été attaqué jusqu'à maintenant (cf. [Minker & Perlis 1986] par exemple), l'a toujours été via des procédés ad'hoc apparemment impossibles à généraliser. Il est abordé dans cet article selon une approche unifiée qui bénéficie en outre d'une grande généralité comme le montre la dernière section. Le premier point est de constater que la propriété de complétude recherchée correspond à ce que les modèles minimaux et les modèles du schéma de circonscription soient les mêmes. Le second point consiste dès lors seulement à comparer la minimisation d'un prédicat dans un modèle telle qu'elle est effectuée par le schéma de circonscription et telle qu'elle est définie par l'ordre sur les modèles. En bref, tout le problème de la complétude de la circonscription tourne autour de l'équivalence, pour la réalisation, dans un modèle quelconque, du prédicat circonscrit, entre minimalité par inclusion ensembliste et minimalité telle que spécifiée au moyen d'une formule logique (en l'occurence toute instance du schéma de circonscription). Cette condition d'équivalence s'exprime alors aisément (c'est l'énoncé du théorème 2) grâce à la définition ci-dessous (tirée de [Chang & Keisler 1973] comme l'essentiel de théorie des modèles nécessaire à cette étude).

Définition 3: Etant donné un modèle $\mathbf{m}$, un sous-ensemble S de $|\mathbf{m}|^n$ est définissable avec paramètres dans $\mathbf{m}$ ssi il existe une formule $F(\vec{x},\vec{y})$ (toutes les variables libres de la formule étant dans $\vec{x},\vec{y}$ avec $\vec{x}$ comportant n variables distinctes) et une séquence $\vec{e}$ (d'éléments de $|\mathbf{m}|$) telles que

$$\mathbf{m} \models F(\vec{s},\vec{e}) \text{ ssi } \vec{s} \in \mathrm{S}$$

Théorème 2: Pour tout modèle $\mathbf{m}$ d'une théorie T, $\mathbf{m}$ est un modèle de $C_{\mathrm{T}}\{P\}$ ssi il n'existe aucun modèle $\mathbf{n}$ de T tel que $\mathbf{n} <_P \mathbf{m}$ où $/P/_{\mathbf{n}}$ serait définissable avec paramètres dans $\mathbf{m}$

Preuve:

(si)

Soit $\mathbf{m}$ un modèle tel qu'indiqué dans l'énoncé du théorème. Soit $F(\vec{x},\vec{y})$ une formule et soit $\vec{e}$ un élément de $|\mathbf{m}|^n$. Supposons que $\mathbf{m} \models \mathrm{T}[F(\,,\vec{e})]$ et $\mathbf{m} \models \{\forall\vec{x}\ F(\vec{x},\vec{e}) \Rightarrow P(\vec{x})\}$. Construisons un modèle $\mathbf{n}$ identique à $\mathbf{m}$ si ce n'est que

$$\mathbf{n} \models P(\vec{s}) \text{ ssi } \mathbf{m} \models F(\vec{s},\vec{e}) \quad \text{pour tout } \vec{s} \in |\mathbf{m}|^k \text{ où } k \text{ est l'arité de } P$$

Alors

$$\mathbf{n} \models \mathrm{T}[P] \text{ ssi } \mathbf{m} \models \mathrm{T}[F(\,,\vec{e})]$$

et donc, $\mathbf{n}$ est un modèle de T. De plus, il est clair que $\mathbf{n} <_P \mathbf{m}$ si $\mathbf{n}$ est distinct de $\mathbf{m}$. Or $|P|_{\mathbf{n}}$ est définissable avec paramètres dans $\mathbf{m}$. Donc, par hypothèse, il est impossible que $\mathbf{n} <_P \mathbf{m}$ et il s'ensuit $\mathbf{n}=\mathbf{m}$. C'est pourquoi $\mathbf{m} \models \{\forall \vec{x}\ P(\vec{x}) \Leftrightarrow F(\vec{x},\vec{e})\}$. Ainsi, $\mathbf{m}$ satisfait l'instance considérée du schéma de circumscription.

(seulement si)

Supposons que $\mathbf{m}$ soit un modèle de $C_T\{P\}$ pour lequel il existe un modèle $\mathbf{n}$ tel que décrit dans l'énoncé du théorème.

Puisque $|P|_{\mathbf{n}}$ est définissable avec paramètres dans $\mathbf{m}$, il existe une formule $F(\vec{x},\vec{y})$ et une séquence $\vec{e}$ d'éléments du domaine de $\mathbf{m}$ telles que

$$\mathbf{m} \models F(\vec{s},\vec{e}) \text{ ssi } \mathbf{n} \models P(\vec{s}) \quad \text{pour tout } \vec{s} \in |\mathbf{m}|^k \text{ où } k \text{ est l'arité de } P$$

Comme $\mathbf{n} <_P \mathbf{m}$, il est clair que $|Q|_{\mathbf{n}} = |Q|_{\mathbf{m}}$ pour tout prédicat Q autre que P

D'où

$$\mathbf{m} \models T[F(\ ,\vec{e})] \text{ ssi } \mathbf{n} \models T[P]$$

Or $\mathbf{n}$ étant un modèle de T, il vient

$$\mathbf{m} \models T[F(\ ,\vec{e})]$$

Etant donné que $\mathbf{n} <_P \mathbf{m}$, il est évident que $|P|_{\mathbf{n}} \subset |P|_{\mathbf{m}}$ et donc

$$\text{si } \mathbf{n} \models P(\vec{s}) \text{ alors } \mathbf{m} \models P(\vec{s}) \quad \text{pour tout } \vec{s} \in |\mathbf{m}|^k \text{ où } k \text{ est l'arité de } P$$

Grâce à ce qui a été établi plus haut comme conséquence du fait que $|P|_{\mathbf{n}}$ est définissable avec paramètres dans $\mathbf{m}$, il s'ensuit

$$\text{si } \mathbf{m} \models F(\vec{s},\vec{e}) \text{ alors } \mathbf{m} \models P(\vec{s}) \quad \text{pour tout } \vec{s} \in |\mathbf{m}|^k \text{ où } k \text{ est l'arité de } P$$

Ce qui donne

$$\mathbf{m} \models \forall \vec{x}\ F(\vec{x},\vec{e}) \Rightarrow P(\vec{x})$$

Sachant que $\mathbf{m}$ est un modèle de $C_T\{P\}$,

$$\mathbf{m} \models \{\forall \vec{x}\ P(\vec{x}) \Rightarrow F(\vec{x},\vec{e})\}$$

Alors, $|P|_{\mathbf{n}} = |P|_{\mathbf{m}}$ et donc $\mathbf{n}=\mathbf{m}$, ce qui contredit les hypothèses.

□

Le théorème 2 permet de retrouver les résultats de complétude partiels connus, en particulier le théorème 3 dont la formulation nécessite deux définitions préalables.

Définition 4: Un symbole de prédicat P est disjonctivement définissable avec paramètres dans une théorie T ssi il existe des formules $F_1, .., F_n$ dans lesquelles P n'apparait pas et telles que

$$T \vdash \exists \vec{y}_1 \forall \vec{x}\, P(\vec{x}) \Leftrightarrow F_1(\vec{x},\vec{y}_1) \vee .. \vee \exists \vec{y}_n \forall \vec{x}\, P(\vec{x}) \Leftrightarrow F_n(\vec{x},\vec{y}_n)$$

Définition 5: Théories bien fondées

Une théorie $\mathbf{T}$ est bien fondée en P ssi tout modèle $\mathbf{m}$ de $\mathbf{T}$ pour lequel il n'existe pas de modèle P-minimal $\mathbf{n}$ de $\mathbf{T}$ tel que $\mathbf{n} <_P \mathbf{m}$, est lui-même un modèle P-minimal de $\mathbf{T}$

Théorème 3: Complétude partielle de la circonscription (Minker & Perlis)

Soit une théorie $\mathbf{T}$ bien fondée en P et telle que P est disjonctivement définissable avec paramètres dans $C_\mathbf{T}\{P\}$. Alors pour toute formule A, $C_\mathbf{T}\{P\} \vdash A$ ssi A est valide dans la classe des modèles P-minimaux de $\mathbf{T}$.

Preuve: Il suffit de montrer que pour une théorie $\mathbf{T}$ remplissant les conditions décrites, les modèles de $C_\mathbf{T}\{P\}$ sont exactement les modèles de $\mathbf{T}$ minimaux relativement à P. De par le théorème 1, il reste seulement à montrer que tous les modèles de $C_\mathbf{T}\{P\}$ sont des modèles de $\mathbf{T}$ minimaux relativement à P.

Supposons qu'il existe un modèle $\mathbf{m}$ de $C_\mathbf{T}\{P\}$ qui ne soit pas un modèle P-minimal de $\mathbf{T}$. Comme $\mathbf{T}$ est bien fondée en P, il en existe un modèle P-minimal $\mathbf{n}$ tel que $\mathbf{n} <_P \mathbf{m}$. Puisque $\mathbf{n}$ est un modèle P-minimal de $\mathbf{T}$, c'est un modèle de $C_\mathbf{T}\{P\}$. D'autre part, P étant disjonctivement définissable avec paramètres dans $C_\mathbf{T}\{P\}$, il existe des formules $F_1, .., F_n$ telles que pour certains $\vec{e}_1, .., \vec{e}_n$

$$\mathbf{n} \models \forall \vec{x}\, P(\vec{x}) \Leftrightarrow F_1(\vec{x}, \vec{e}_1) \vee .. \vee \forall \vec{x}\, P(\vec{x}) \Leftrightarrow F_n(\vec{x}, \vec{e}_n)$$

Alors, pour un certain i,

$$\mathbf{n} \models \forall \vec{x}\, P(\vec{x}) \Leftrightarrow F_i(\vec{x}, \vec{e}_i)$$

où le symbole P n'apparait pas dans $F_i(\vec{x}, \vec{e}_i)$.

Ce qui fait que $F_i(\ , \vec{e}_i)$ a la même réalisation dans $\mathbf{m}$ et dans $\mathbf{n}$ car les deux modèles sont comparables par $<_P$. Ainsi, $|P|_\mathbf{n}$ est définissable avec paramètres dans $\mathbf{m}$ et comme $\mathbf{n}$ est un modèle de $\mathbf{T}$ tel que $\mathbf{n} <_P \mathbf{m}$, de par le théorème 2, $\mathbf{m}$ n'est pas un modèle de $C_\mathbf{T}\{P\}$, ce qui contredit les hypothèses.

□

Le théorème 3 implique que les théories n'ayant que des modèles finis forment une classe sur laquelle la circonscription est complète. C'est pourquoi le théorème 4 qui établit l'incomplétude de la circonscription sur la classe des théories universelles finitaires (c'est-à-dire le langage de la forme skolémisée d'une telle théorie ne comporte pas de symbole de fonction d'arité non nulle) constitue un résultat très fort qui contient les théorèmes d'incomplétude obtenus à ce jour pour la circonscription.

Théorème 4: Incomplétude de la circonscription

Il existe une théorie universelle finitaire sur laquelle la circonscription est incomplète.

Preuve:

Soit la théorie $\mathbf{T}$ constituée des cinq axiomes ci-dessous.

$$I(0)$$
$$\forall x \forall y \ I(x) \wedge (S(x,y) \vee S(y,x)) \Rightarrow I(y)$$
$$\forall x \ \neg S(x,0)$$
$$\forall x \forall y \forall z \ S(x,z) \wedge S(y,z) \Rightarrow x = y$$
$$\forall x \forall y \forall z \ S(x,y) \wedge S(x,z) \Rightarrow z = y$$

On construit un modèle $\mathbf{m}$ de cette théorie de sorte que le domaine de $\mathbf{m}$ soit composé de deux copies des entiers naturels $\mathbb{N} + \mathbb{N}'$ et que $\mathbf{m}$ interprète S comme la relation correspondant au successeur. En exigeant de plus $\mathbf{m} \models \forall x \ I(x)$, on obtient un modèle qui n'est pas minimal relativement à I. De par le second axiome, il n'y a qu'un seul modèle $\mathbf{n}$ de $\mathbf{T}$ tel que $\mathbf{n} <_I \mathbf{m}$: c'est le modèle dans lequel I n'est satisfait que sur $\mathbb{N}$. Cependant, le théorème 2 montre que $\mathbf{m}$ est un modèle du schéma de circumscription de I dans $\mathbf{T}$ parce que $\mathbb{N}$ n'est pas définissable (même avec paramètres) dans $\mathbf{m}$. En effet, il est facile de vérifier que la théorie $\mathbf{T_m}$ induite par $\mathbf{m}$ satisfait aux conditions de la proposition 1 ci-dessous (parce que si $b \notin |\mathbf{m}|$ alors un plus petit $\mathbf{T_m}$-sur-modèle $\mathbf{m}(b)$ contenant b est isomorphe à $\mathbb{N} + \mathbb{N}' + \mathbb{Z}''$, i.e. $\mathbf{m}$ étendu par une copie de $\mathbb{Z}$). En conséquence, $\mathbf{m}$ est tel que ses seules parties définissables (y compris avec paramètres) sont soit finies soit cofinies.

L'incomplétude est due au fait que $\mathbf{m}$ falsifie $\forall x \ I(x) \Rightarrow (x = 0 \vee \exists y S(y,x))$, qui est une formule satisfaite dans tout modèle de $\mathbf{T}$ minimal relativement à I.

□

Que les seules parties définissables (avec paramètres) d'un modèle soient ou finies ou cofinies est un résultat très utile dans toute cette étude et pour lequel un critère assez général est décrit ci-après, à la suite de quelques définitions.

Définition 6: Etant donné deux modèles $\mathbf{m}$ et $\mathbf{m}'$ d'une théorie $\mathbf{T}$, $\mathbf{m}$ est un $\mathbf{T}$-sous-modèle de $\mathbf{m}'$ et $\mathbf{m}'$ est un $\mathbf{T}$-sur-modèle de $\mathbf{m}$ ssi $|\mathbf{m}| \subseteq |\mathbf{m}'|$ et sur $|\mathbf{m}|$ chaque symbole du langage (constante, fonction, prédicat) est interprété de la même façon par chacun des deux modèles.

Définition 7: Théories à sur-modèles minimaux uniformes

Une théorie à sur-modèles minimaux uniformes est une théorie T satisfaisant aux deux conditions suivantes:

i. Si $\mathbf{m}'$ est un T-sur-modèle de $\mathbf{m}$ alors
   pour tout $b \in |\mathbf{m}'|-|\mathbf{m}|$
   il existe un plus petit T-sous-modèle de $\mathbf{m}'$, noté $\mathbf{m}(b)$, contenant $|\mathbf{m}| \cup \{b\}$
   (l'ordre étant celui de l'inclusion ensembliste sur le domaine des modèles)

ii. Si $\mathbf{m}'$ et $\mathbf{m}''$ sont deux T-sur-modèles de $\mathbf{m}$ alors
   pour tous b', b'' où $b' \in |\mathbf{m}'|-|\mathbf{m}|$ et $b'' \in |\mathbf{m}''|-|\mathbf{m}|$
   $\mathbf{m}(b')$ et $\mathbf{m}(b'')$ sont isomorphes par un isomorphisme laissant invariants les éléments de $|\mathbf{m}|$ et envoyant b' sur b''

Définition 8: Une théorie T est modèle-complète ssi pour tout modèle $\mathbf{m}$ de T la théorie $\mathrm{T} \cup \mathrm{D}_{\mathbf{m}}$ est complète (où le diagramme $\mathrm{D}_{\mathbf{m}}$ de $\mathbf{m}$ est constitué de tous les littéraux du langage de $\mathbf{m}$ satisfaits dans $\mathbf{m}$).

Lemme 1: Toute théorie à sur-modèles minimaux uniformes est modèle-complète

Preuve: Etant donné une théorie T, on montre par récurrence sur n que si $\mathbf{m}'$ est un T-sur-modèle de $\mathbf{m}$ et A est une formule à n quantificateurs alors $\mathbf{m} \models A$ ssi $\mathbf{m}' \models A$.

Le résultat pour le cas sans quantificateurs est connu, aussi supposons la propriété vraie pour n et prenons pour A la formule $\exists x B(x)$ où B a n quantificateurs.

Nous nous limitons à la démonstration du cas des modèles infinis.

Supposons $\mathbf{m}' \models A$. Alors $\mathbf{m}' \models B(b)$ pour un certain b de $|\mathbf{m}'|$. Comme $\mathbf{m}'$ est un T-sur-modèle de $\mathbf{m}(b)$, on a $\mathbf{m}(b) \models B(b)$. Posons $\mathrm{H}_b = \{\neg b = m \;/\; m \in |\mathbf{m}|\}$. Soit $\mathbf{m}''$ un modèle de $\mathrm{T} \cup \mathrm{D}_{\mathbf{m}} \cup \mathrm{H}_b$. Sachant que si $\mathbf{m}$ est un modèle d'une théorie T alors tout modèle de $\mathrm{T} \cup \mathrm{D}_{\mathbf{m}}$ est isomorphe à un T-sur-modèle de $\mathbf{m}$, il existe donc un modèle isomorphe à $\mathbf{m}''$ qui est un T-sur-modèle de $\mathbf{m}$ et qui par les propriétés i. et ii. satisfait $\exists x B(x)$ car $\mathbf{m}(b) \models B(b)$. D'où $\mathbf{m}'' \models \exists x B(x)$. Ceci vaut pour tout $\mathbf{m}''$, et il vient $\mathrm{T} \cup \mathrm{D}_{\mathbf{m}} \cup \mathrm{H}_b \vdash \exists x B(x)$. Par compacité il existe une conjonction finie $C(b)$ d'éléments de H_b telle que $\mathrm{T} \cup \mathrm{D}_{\mathbf{m}} \cup \{C(b)\} \vdash \exists x B(x)$. Il suit $\mathrm{T} \cup \mathrm{D}_{\mathbf{m}} \vdash C(b) \Rightarrow A$. D'où $\mathrm{T} \cup \mathrm{D}_{\mathbf{m}} \vdash \exists x C(x) \Rightarrow A$ puisque b n'apparait ni dans A ni dans $\mathrm{T} \cup \mathrm{D}_{\mathbf{m}}$. Comme $\mathbf{m}$ est infini par hypothèse, $\mathbf{m} \models \exists x C(x)$. Donc $\mathbf{m} \models A$.

Supposons $\mathbf{m} \models A$. Dans ce cas, $\mathbf{m} \models B(b)$ pour un certain b de $|\mathbf{m}|$. Il est alors évident que $\mathbf{m}' \models B(b)$ et $\mathbf{m}' \models A$.

Si on prend pour A la formule $\forall x B(x)$ on est ramené aux cas précédents car $\neg A$ est alors la formule $\exists x \neg B(x)$.

□

Proposition 1: Si $\mathbf{m}$ est un modèle d'une théorie $\mathbf{T}$ à sur-modèles minimaux uniformes et si F est une formule du langage de $\mathbf{m}$ à une variable libre alors l'ensemble des éléments de $|\mathbf{m}|$ qui satisfont F est fini ou cofini

Preuve: Supposons que F contredise l'énoncé. Dans ce cas, $\Omega_1 = \mathbf{T} \cup \mathbf{D_m} \cup \mathbf{H}_b \cup \{F(b)\}$ et $\Omega_2 = \mathbf{T} \cup \mathbf{D_m} \cup \mathbf{H}_b \cup \{\neg F(b)\}$ sont finiment satisfiables: toute partie finie de $\mathbf{H}_b$ ne distingue b que d'un nombre fini d'éléments de $|\mathbf{m}|$ et donc (l'élément interprétant) b peut aussi bien être choisi dans l'ensemble des éléments restants qui satisfont F que dans l'ensemble des éléments restants qui satisfont $\neg F$, aucun des deux n'étant vide à partir du moment où la réalisation de F dans $\mathbf{m}$ n'est ni finie ni cofinie. Par compacité, Ω_1 et Ω_2 sont satisfiables.
Il est connu que si $\mathbf{m}$ est un modèle d'une théorie $\mathbf{T}$ alors tout modèle de $\mathbf{T} \cup \mathbf{D_m}$ est isomorphe à un $\mathbf{T}$-sur-modèle de $\mathbf{m}$. Il existe donc un modèle $\mathbf{m}_1$ de Ω_1 et un modèle $\mathbf{m}_2$ de Ω_2 qui sont des $\mathbf{T}$-sur-modèles de $\mathbf{m}$. Soit $b_1 \in |\mathbf{m}_1|-|\mathbf{m}|$ tel que $\mathbf{m}_1 \models F(b_1)$ et soit $b_2 \in |\mathbf{m}_2|-|\mathbf{m}|$ tel que $\mathbf{m}_2 \models \neg F(b_2)$. Or, si $\mathbf{n}'$ est un $\mathbf{T}$-sur-modèle de $\mathbf{n}$ où $\mathbf{T}$ est une théorie modèle-complète alors $\mathbf{n}' \models \alpha(e_1,..,e_n)$ ssi $\mathbf{n} \models \alpha(e_1,..,e_n)$ pour toute formule $\alpha(x_1,..,x_n)$ à n variables libres et toute suite $e_1,..,e_n$ d'éléments de $|\mathbf{n}|$. Donc le lemme 1 entraine à la fois $\mathbf{m}(b_1) \models F(b_1)$ et $\mathbf{m}(b_2) \models \neg F(b_2)$, ce qui est absurde puisqu'ils doivent être isomorphes de par l'hypothèse ii.

□

Les théorèmes 3 et 4 montrent la puissance du théorème 2, qui sert aussi bien à obtenir des résultats de complétude que d'incomplétude pour la circonscription.

La prochaine section présente différentes versions du théorème 2, adaptées à diverses variétés de circonscriptions.

3 Généralisation et restrictions: variétés de circonscriptions

Dans la circonscription telle qu'elle a été décrite jusqu'ici, l'idée est de minimiser P, les autres prédicats étant fixés. Ceci est relativement limitatif: par exemple, la circonscription de P dans la théorie ayant pour unique axiome $\neg P(a) \Rightarrow Q(b)$ ne permet pas de dériver $\neg P(a)$ car autrement $Q(b)$ pourrait aussi être dérivé. Il est donc intéressant d'étudier ce qui passe lorsque qu'un prédicat autre que le prédicat circonscrit est autorisé à varier [McCarthy 1986]. C'est-à-dire que si P est le prédicat circonscrit et Q le prédicat autorisé à varier, l'idée est de minimiser P avec tous les autres prédicats, sauf éventuellement Q, fixés.

Définition 1': Circonscription standard

Le schéma pour la circonscription du prédicat P, avec le prédicat Q autorisé à varier, dans la théorie finiment axiomatisable T est de la forme

$$\{\mathrm{T}[\phi_P(\,,\vec{y}), \phi_Q(\,,\vec{z})] \ \wedge\ \forall\vec{x}\,(\phi_P(\vec{x},\vec{y}) \Rightarrow P(\vec{x}))\} \ \Rightarrow\ \{\forall\vec{x}\,(P(\vec{x}) \Rightarrow \phi_P(\vec{x},\vec{y}))\}$$

où
- $\mathrm{T}[\phi_P(\,,\vec{y}), \phi_Q(\,,\vec{z})]$ résulte du remplacement (sur la base $P(\vec{x}) \Leftrightarrow \phi_P(\vec{x},\vec{y})$ et $Q(\vec{u}) \Leftrightarrow \phi_Q(\vec{u},\vec{z})$) de P et Q par $\phi_P(\,,\vec{y})$ et $\phi_Q(\,,\vec{z})$ respectivement partout dans $\mathrm{T}[P, Q]$ (qui est une axiomatisation finie de T)
- $\phi_P(\vec{x},\vec{y})$ ainsi que $\phi_Q(\vec{u},\vec{z})$ peut être toute formule dont les variables libres $\vec{y}$ et $\vec{z}$ respectivement ne doivent pas se retrouver liées dans $\mathrm{T}[\phi'_P(\,,\vec{y}), \phi_Q(\,,\vec{z})]$

Définition 2': Un modèle $\mathbf{m}$ d'une théorie T est un modèle minimal de T relativement à P avec Q autorisé à varier ssi

il n'existe aucun modèle $\mathbf{m}'$ de T tel que $\mathbf{m}' <_{P/Q} \mathbf{m}$

où $\mathbf{n} <_{P/Q} \mathbf{m}$ ssi $|P|_{\mathbf{n}} \subset |P|_{\mathbf{m}}$ avec $\mathbf{m}$ et $\mathbf{n}$ identiques par ailleurs sauf peut-être pour Q

Pour adapter le théorème 2 à la circonscription standard de la définition 1' et sa théorie des modèles minimaux de la définition 2', il suffit de tenir compte de la définissabilité du prédicat qui est autorisé à varier et la preuve donnée s'applique moyennant ce point précis.

Théorème 2': Pour tout modèle $\mathbf{m}$ d'une théorie T, $\mathbf{m}$ est un modèle de $C_{\mathrm{T}}\{P, Q\}$ ssi il n'existe aucun modèle $\mathbf{n}$ de T tel que $\mathbf{n} <_{P/Q} \mathbf{m}$ où $|P|_{\mathbf{n}}$ et $|Q|_{\mathbf{n}}$ seraient définissables avec paramètres dans $\mathbf{m}$

Les résultats de consistance, de complétude partielle et d'incomplétude sont également valables pour cette circonscription généralisée.

Des problèmes d'incohérence [Etherington, Mercer & Reiter 1985] survenant dans le cas de la circonscription de prédicats avec définition récursive, une restriction de la circonscription standard a été avancée récemment [Mott 1987] pour préserver la cohérence des théories sur lesquelles la circonscription opère. La proposition de Mott s'exprime comme suit.

Définition 1″: Circonscription non récursive

Le schéma pour la circonscription du prédicat P, avec le prédicat Q autorisé à varier, dans la théorie finiment axiomatisable $\mathbf{T}$ est de la forme

$$\{\mathbf{T}[\phi_P(\,,\vec{y}),\phi_Q(\,,\vec{z})] \wedge \forall\vec{x}(\phi_P(\vec{x},\vec{y})\Rightarrow P(\vec{x}))\} \Rightarrow \{\forall\vec{x}(P(\vec{x})\Rightarrow\phi_P(\vec{x},\vec{y}))\}$$

où
- $\mathbf{T}[\phi_P(\,,\vec{y}),\phi_Q(\,,\vec{z})]$ résulte du remplacement (sur la base $P(\vec{x})\Leftrightarrow\phi_P(\vec{x},\vec{y})$ et $Q(\vec{u})\Leftrightarrow\phi_Q(\vec{u},\vec{z})$) de P et Q par $\phi_P(\,,\vec{y})$ et $\phi_Q(\,,\vec{z})$ respectivement partout dans $\mathbf{T}[P,Q]$ (qui est une axiomatisation finie de $\mathbf{T}$)
- $\phi_P(\vec{x},\vec{y})$ ainsi que $\phi_Q(\vec{u},\vec{z})$ peut être toute formule dans laquelle ni P ni Q n'apparaissent et dont les variables libres $\vec{y}$ et $\vec{z}$ respectivement ne doivent pas se retrouver liées dans $\mathbf{T}[\phi_P(\,,\vec{y}),\phi_Q(\,,\vec{z})]$

Pour adapter le théorème 2 à la circonscription non récursive de la définition 1″ et sa théorie des modèles minimaux de la définition 2′, il s'agit cette fois de mettre en jeu la définissabilité des prédicats dans un sous-langage du langage de la théorie.

Théorème 2″: Pour tout modèle $\mathbf{m}$ d'une théorie $\mathbf{T}$, $\mathbf{m}$ est un modèle de la version non récursive de $C_{\mathbf{T}}\{P,Q\}$ ssi il n'existe aucun modèle $\mathbf{n}$ de $\mathbf{T}$ tel que $\mathbf{n} <_{P/Q} \mathbf{m}$ où $|P|_{\mathbf{n}}$ et $|Q|_{\mathbf{n}}$ seraient définissables avec paramètres dans $\mathbf{m}$ avec le langage obtenu en retirant les symboles P et Q

Mott a également proposé de se restreindre à un schéma de circonscription dont toutes les instances seraient des formules closes.

Définition 1''': Circonscription close

Le schéma pour la circonscription du prédicat P, avec le prédicat Q autorisé à varier, dans la théorie finiment axiomatisable T est de la forme

$$\{\mathrm{T}[\phi_P, \phi_Q] \wedge \forall\vec{x}(\phi_P(\vec{x}) \Rightarrow P(\vec{x}))\} \Rightarrow \{\forall\vec{x}(P(\vec{x}) \Rightarrow \phi_P(\vec{x}))\}$$

où
- $\mathrm{T}[\phi_P, \phi_Q]$ résulte du remplacement (sur la base $P(\vec{x}) \Leftrightarrow \phi_P(\vec{x})$ et $Q(\vec{u}) \Leftrightarrow \phi_Q(\vec{u})$) de P et Q par ϕ_P et ϕ_Q respectivement partout dans $\mathrm{T}[P, Q]$ (qui est une axiomatisation finie de T)
- $\phi_P(\vec{x})$ a $\vec{x}$ pour seules variables libres tandis que si ϕ_Q a des variables libres, elles se retrouvent toutes liées dans $\mathrm{T}[\phi_P, \phi_Q]$

Pour adapter le théorème 2 à la circonscription close de la définition 1''' et sa théorie des modèles minimaux de la définition 2', il convient de considérer la définissabilité des prédicats à l'aide de formules n'ayant d'autres variables libres que les variables définissantes.

Théorème 2''': Pour tout modèle $\mathbf{m}$ d'une théorie T, $\mathbf{m}$ est un modèle de la version close de $C_{\mathrm{T}}\{P, Q\}$ ssi il n'existe aucun modèle $\mathbf{n}$ de T tel que $\mathbf{n} <_{P/Q} \mathbf{m}$ où $|P|_{\mathbf{n}}$ et $|Q|_{\mathbf{n}}$ seraient définissables (sans paramètres) dans $\mathbf{m}$

Finalement, toutes ces variantes permettent une formulation unique.

Théorème 5: Pour tout modèle $\mathbf{m}$ d'une théorie T, $\mathbf{m}$ est un modèle du $\mathrm{F}_{P/Q}$-schéma de circonscription $C_{\mathrm{T}}\{P, Q\}$ ssi il n'existe aucun modèle $\mathbf{n}$ de T tel que $\mathbf{n} <_{P/Q} \mathbf{m}$ où $|P|_{\mathbf{n}}$ et $|Q|_{\mathbf{n}}$ seraient $\mathrm{F}_{P/Q}$-définissables dans $\mathbf{m}$

Dans cet énoncé, $\mathrm{F}_{P/Q}$ dénote

-dans le cas de la circonscription standard, l'ensemble de toutes les formules du langage

-dans le cas de la circonscription non récursive, l'ensemble de toutes les formules du langage obtenu en retirant le symbole du prédicat circonscrit et le symbole du prédicat autorisé à varier

-dans le cas de la circonscription close, l'ensemble de toutes les formules ayant un nombre de variables libres ne dépassant pas l'arité du prédicat concerné (d'une part, le prédicat circonscrit, et d'autre part, le prédicat autorisé à varier)

A chaque fois il est bien entendu que la substitution d'une formule donnée de $\mathrm{F}_{P/Q}$ dans un schéma de circonscription repose sur un choix raisonnable des variables du schéma.

Ce théorème peut servir bien évidemment à établir des résultats de complétude partiels ou d'incomplétude pour les circonscriptions concernées, mais aussi à démontrer que les versions restrictives peuvent se révéler incomplètes là où la circonscription standard est complète [Besnard, Mercer & Moinard 1987] comme l'illustre la théorie de l'exemple 2.

Exemple 2: Intuitivement, la minimisation d'une relation binaire totale doit aboutir à une relation totale antisymmétrique. Cependant, le schéma de circonscription de P dans la théorie $\mathbf{T}_2$ ayant pour seul axiome

$$\forall x \forall y \; P(x,y) \vee P(y,x)$$

permet [Moinard 1988] de dériver la conclusion souhaitée

$$\forall x \forall y \; P(x,y) \wedge P(y,x) \Rightarrow x = y$$

pour une instance qui est à la fois récursive et ouverte,

$$\phi_P(x,y,z,t) \text{ valant } \neg P(y,x) \vee \{ P(x,y) \wedge P(y,x) \wedge (\neg x = z \vee \neg y = t) \}$$

Il est plutôt facile de montrer qu'appliquée à la théorie $\mathbf{T}_2$ la version close de $C_{\mathbf{T}}\{P\}$ ne permet pas de déduire l'antisymmétrie de P. En se restreignant à un langage ayant P pour seul symbole non logique, il n'est que de considérer le modèle $\mathbf{m}$ de domaine $\{a,b\}$ (avec $a \neq b$) tel que $/P/_{\mathbf{m}} = |\mathbf{m}|^2$. Alors $\mathbf{m} \models F(a,b) \Leftrightarrow F(b,a)$ pour toute formule F du langage de $\mathbf{m}$. En effet, l'ensemble des formules de ce langage qui ont cette propriété contient les formules atomiques et est stable pour $\neg$, $\wedge$ et $\forall$. Il suffit maintenant d'appliquer le théorème 2''' et par complétude de la logique du premier ordre il est clair que la formule d'antisymmétrie pour P ne peut être dérivée puisqu'elle est falsifiée par $\mathbf{m}$.

Pour la même théorie $\mathbf{T}_2$, la version non récursive de $C_{\mathbf{T}}\{P\}$ ne permet pas de déduire la formule $\forall x \forall y \; P(x,y) \wedge P(y,x) \Rightarrow x = y$. Pour le confirmer, il n'y a qu'à exhiber le modèle $\mathbf{m}$ ayant pour domaine $\mathbb{N}$ et tel que $/P/_{\mathbf{m}} = \{(1,2n+1)\} \cup \{(2n,1)\} \cup \{(n,m) \;/\; n \neq 1 \text{ et } m \neq 1\}$. Pour le langage ayant " = " pour seul symbole de prédicat, la théorie induite par $\mathbf{m}$ est une théorie à sur-modèles minimaux uniformes étant donné que pour tout $b \notin |\mathbf{m}|$ il est possible de définir un modèle $\mathbf{m}(b)$ de domaine $|\mathbf{m}| \cup \{b\}$ et ayant les propriétés appropriées. Supposons maintenant qu'il existe une partie $\mathbf{P}'$ de $/P/_{\mathbf{m}}$ satisfaisant l'unique axiome de la théorie et qui soit définissable (avec paramètres – qui sont des éléments de $|\mathbf{m}|$) par une formule $F'(u,v)$ du langage ayant " = " pour seul symbole de prédicat. Appelons $F(x)$ la formule $F'(1,x)$. La proposition 1 étant appliquée, $/F/_{\mathbf{m}}$ est fini ou cofini. Mais $(1,2n) \notin /P/_{\mathbf{m}}$ et donc $(1,2n) \notin \mathbf{P}'$ i.e. $2n \notin /F/_{\mathbf{m}}$. Ainsi $/F/_{\mathbf{m}}$ n'est pas cofini. Par suite, $/F/_{\mathbf{m}}$ est fini. Alors il existe n tel que $2n+1 \notin /F/_{\mathbf{m}}$ c'est-à-dire $(1, 2n+1) \notin \mathbf{P}'$. Or $(2n+1, 1) \notin /P/_{\mathbf{m}}$. D'où $(2n+1, 1) \notin \mathbf{P}'$, ce qui contredit l'hypothèse que $\mathbf{P}'$ satisfait l'axiome de la théorie. Le théorème 2″ montre que $\mathbf{m}$ est un modèle de la version non récursive de la circonscription de P dans T_2 et donc la formule $\forall x \forall y\ P(x,y) \wedge P(y,x) \Rightarrow x = y$ n'étant pas satisfaite par $\mathbf{m}$ ne peut se déduire de cette version de $C_{\mathrm{T}}\{P\}$.

Finalement, le théorème 2 (et similairement le théorème 5) s'inscrit parfaitement dans l'univers des circonscriptions à un autre titre encore, ayant trait à la circonscription exprimée en logique du second ordre.

Définition 9: Circonscription du second ordre

L'axiome du second ordre pour la circonscription du prédicat P, avec le prédicat Q autorisé à varier, dans la théorie finiment axiomatisable T, s'écrit

$$\forall U \forall V\ \{\mathrm{T}[U, V]\ \wedge\ \forall \vec{x}\,(U(\vec{x}) \Rightarrow P(\vec{x}))\}\ \Rightarrow\ \{\forall \vec{x}\,(P(\vec{x}) \Rightarrow U(\vec{x}))\}$$

Ici les formules ϕ_P, ϕ_Q du schéma d'axiomes du premier ordre sont remplacées par les variables du second ordre U, V. Il est clair qu'avec la circonscription du second ordre, les modèles minimaux et les modèles standard du second ordre sont les mêmes. C'est pourquoi la circonscription du second ordre hérite des résultats de consistance et d'incomplétude de la logique du second ordre. Ce que le théorème 2 illustre bien est la différence de puissance entre le schéma de circonscription (qui est du premier ordre) et l'axiome de circonscription (qui est du second ordre): dans le cas du schéma, la réalisation du prédicat circonscrit (supposé unaire pour simplifier) est minimisée relativement à l'ensemble des parties définissables du domaine du modèle alors que dans le cas de l'axiome, la réalisation du prédicat circonscrit est minimisée relativement à l'ensemble de toutes les parties du domaine du modèle. Retournant à la théorie donnée dans la preuve du théorème 4, il est intéressant de constater que l'axiome de circonscription du second ordre permet de dériver la formule

$$\forall x\ I(x) \Rightarrow (x = 0 \vee \exists y S(y, x))$$

qui est hors de portée du schéma de circonscription du premier ordre. Avec l'axiome de circonscription du second ordre

$$\forall U\ \mathrm{T}[U] \wedge \forall x U(x) \Rightarrow I(x)\ \ \Rightarrow\ \ (\forall x I(x) \Rightarrow U(x))$$

il suffit [Moinard 1988] de particulariser $U(x)$ en

$$\forall Z\ \ (Z(0) \wedge \forall t \forall u\ Z(t) \wedge S(t,u) \Rightarrow Z(u)) \Rightarrow Z(x)$$

4 Références

[Besnard Ph., Mercer R. E. & Moinard Y. 1987] *Some Limitations of Closed and Nonrecursive Circumscription*, à paraitre.

[Chang C. C. & Keisler H. J. 1973] *Model Theory*, North-Holland, Amsterdam.

[Etherington D. W., Mercer R. E. & Reiter R. 1985] *On the Adequacy of Predicate Circumscription for Closed World Reasoning*, Computational Intelligence 1, pp. 11-15.

[McCarthy J. 1980] *Circumscription - A Form of Non-Monotonic Reasoning*, Artificial Intelligence 13, pp. 27-39.

[McCarthy J. 1986] *Applications of Circumscription to Formalizing Commonsense Knowledge*, Artificial Intelligence 28, pp. 89-116.

[Minker J. & Perlis D. 1986] *Completeness Results for Circumscription*, Artificial Intelligence 28, pp. 29-42.

[Moinard Y. 1988] *Contribution à l'étude de la circonscription*, Thèse, Université de Rennes I, en préparation.

[Mott P. L. 1987] *A Theorem on the Consistency of Circumscription*, Artificial Intelligence 31, pp. 87-98.

Imprimé en France
par
l'Institut National de Recherche en Informatique et en Automatique