

Fixpoint logics with tree-like models

Michael Vanden Boom

University of Oxford

Oxford Information Systems Seminar
November 2015

Including joint work with
Michael Benedikt, Pierre Bourhis, Balder ten Cate, and Thomas Colcombet

Fixpoint logics

Fixpoint logics give mechanism to express
dynamic, recursive properties.

Example

binary relation R , unary relation P

“from x , it is possible to R -reach some P -element”

$$[\text{Reach-}P](x) := [\text{Ifp}_{Y,y} . \exists z (Ryz \wedge (Pz \vee Yz))](x)$$

Least fixpoint

Consider $\psi(\mathbf{y}, Y)$ positive in Y (of arity $m = |\mathbf{y}|$).

For all structures \mathfrak{A} , the formula ψ induces a monotone operation

$$\begin{aligned} \mathcal{P}(A^m) &\longrightarrow \mathcal{P}(A^m) \\ V &\longmapsto \psi_{\mathfrak{A}}(V) := \{\mathbf{a} \in A^m : \mathfrak{A}, \mathbf{a}, V \models \psi\} \end{aligned}$$

\Rightarrow there is a unique **least fixpoint** $[\text{lfp}_{Y, \mathbf{y}} \psi(\mathbf{y}, Y)]_{\mathfrak{A}} := \bigcup_{\alpha} \psi_{\mathfrak{A}}^{\alpha}$

$$\begin{aligned} \psi_{\mathfrak{A}}^0 &:= \emptyset \\ \psi_{\mathfrak{A}}^{\alpha+1} &:= \psi_{\mathfrak{A}}(\psi_{\mathfrak{A}}^{\alpha}) \\ \psi_{\mathfrak{A}}^{\lambda} &:= \bigcup_{\alpha < \lambda} \psi_{\mathfrak{A}}^{\alpha} \end{aligned}$$

Least fixpoint

Consider $\psi(\mathbf{y}, Y)$ positive in Y (of arity $m = |\mathbf{y}|$).

For all structures \mathfrak{A} , the formula ψ induces a monotone operation

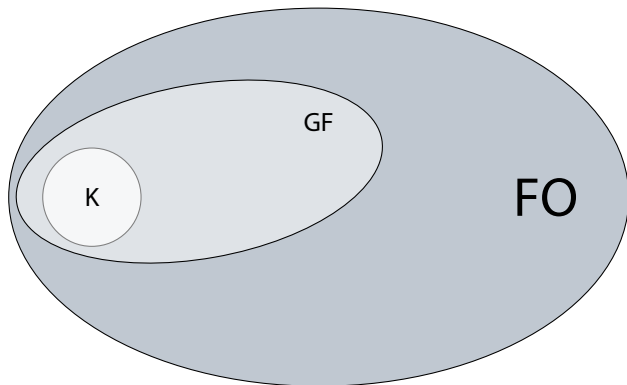
$$\begin{aligned} \mathcal{P}(A^m) &\longrightarrow \mathcal{P}(A^m) \\ V &\longmapsto \psi_{\mathfrak{A}}(V) := \{\mathbf{a} \in A^m : \mathfrak{A}, \mathbf{a}, V \models \psi\} \end{aligned}$$

\Rightarrow there is a unique **least fixpoint** $[\mathbf{lfp}_{Y,y}.\psi(\mathbf{y}, Y)]_{\mathfrak{A}} := \bigcup_{\alpha} \psi_{\mathfrak{A}}^{\alpha}$

$$\begin{aligned} \psi_{\mathfrak{A}}^0 &:= \emptyset \\ \psi_{\mathfrak{A}}^{\alpha+1} &:= \psi_{\mathfrak{A}}(\psi_{\mathfrak{A}}^{\alpha}) \\ \psi_{\mathfrak{A}}^{\lambda} &:= \bigcup_{\alpha < \lambda} \psi_{\mathfrak{A}}^{\alpha} \end{aligned}$$

Semantics of fixpoint operator: $\mathfrak{A}, \mathbf{a} \models [\mathbf{lfp}_{Y,y}.\psi(\mathbf{y}, Y)](\mathbf{x})$ iff $\mathbf{a} \in \bigcup_{\alpha} \psi_{\mathfrak{A}}^{\alpha}$

Some decidable fragments of first-order logic

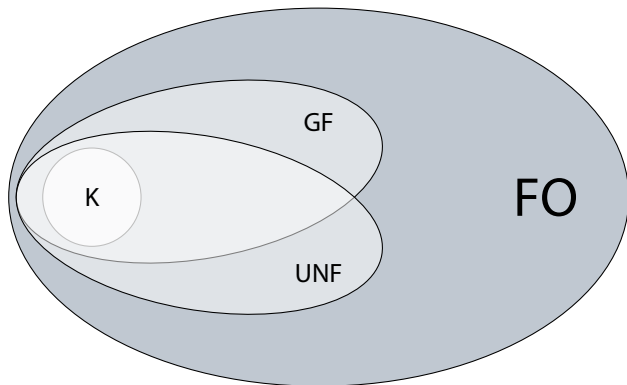


constrain
quantification

$$\exists x(G(xy) \wedge \psi(xy))$$
$$\forall x(G(xy) \rightarrow \psi(xy))$$

[Andréka, van Benthem,
Németi '95-'98]

Some decidable fragments of first-order logic



constrain
quantification

$$\begin{aligned} &\exists x(G(xy) \wedge \psi(xy)) \\ &\forall x(G(xy) \rightarrow \psi(xy)) \end{aligned}$$

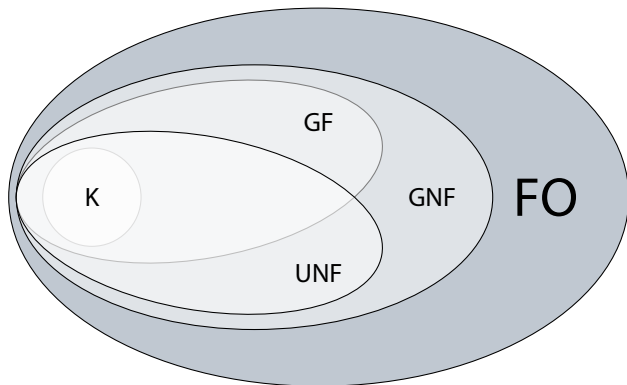
[Andréka, van Benthem,
Németi '95-'98]

constrain
negation

$$\begin{aligned} &\exists x(\psi(xy)) \\ &\neg\psi(x) \end{aligned}$$

[ten Cate, Segoufin '11]

Some decidable fragments of first-order logic



constrain
quantification

$$\exists x(G(xy) \wedge \psi(xy))$$
$$\forall x(G(xy) \rightarrow \psi(xy))$$

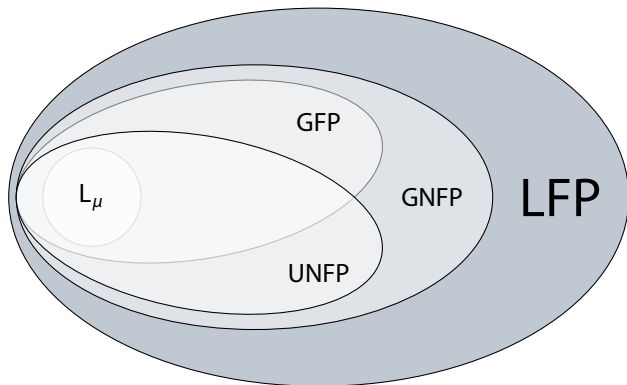
[Andréka, van Benthem,
Németi '95-'98]

constrain
negation

$$\exists x(\psi(xy))$$
$$G(xy) \wedge \neg\psi(xy)$$

[ten Cate, Segoufin '11]
[Bárány, ten Cate, Segoufin '11]

Some decidable fragments of LFP (fixpoint extension of FO)



constrain
quantification

$$\exists x(G(xy) \wedge \psi(xy))$$
$$\forall x(G(xy) \rightarrow \psi(xy))$$

[Andréka, van Benthem,
Németi '95-'98]

constrain
negation

$$\exists x(\psi(xy))$$
$$G(xy) \wedge \neg\psi(xy)$$

[ten Cate, Segoufin '11]

[Bárány, ten Cate, Segoufin '11]

Guarded fixpoints: tuples in fixpoint are guarded by atom in original signature.
(UNFP has only monadic fixpoints, which are trivially guarded.)

Guarded negation fixpoint logic (GNFP)

Fix some signature σ of relations and constants.

Syntax of GNFP[σ]

$$\varphi ::= R\mathbf{t} \mid Y\mathbf{t} \mid \varphi \wedge \varphi \mid \varphi \vee \varphi \mid \exists y(\psi(\mathbf{x}y)) \mid G(\mathbf{x}) \wedge \neg\psi(\mathbf{x}) \mid$$
$$[\text{Ifp}_{Y,y} . G(\mathbf{y}) \wedge \varphi(\mathbf{y}, Y, \mathbf{Z})](\mathbf{t}) \quad \text{where } Y \text{ only occurs positively in } \varphi$$

where R and G are relations in σ or $=$, and \mathbf{t} is a tuple over variables and constants.

Guarded negation fixpoint logic (GNFP)

Fix some signature σ of relations and constants.

Syntax of GNFP[σ]

$$\varphi ::= R\mathbf{t} \mid Y\mathbf{t} \mid \varphi \wedge \varphi \mid \varphi \vee \varphi \mid \exists y(\psi(\mathbf{x}y)) \mid G(\mathbf{x}) \wedge \neg\psi(\mathbf{x}) \mid \\ [\text{Ifp}_{Y,y} . G(\mathbf{y}) \wedge \varphi(\mathbf{y}, Y, \mathbf{Z})](\mathbf{t}) \quad \text{where } Y \text{ only occurs positively in } \varphi$$

where R and G are relations in σ or $=$, and \mathbf{t} is a tuple over variables and constants.

Examples

- unions of conjunctive queries (in GNF)
- frontier-guarded tgds (in GNF):
$$\forall xyz ((Rxy \wedge Ryz) \rightarrow \exists w(Twyz)) \equiv \neg \exists xyz (Rxy \wedge Ryz \wedge \neg \exists w(Twyz))$$

Guarded negation fixpoint logic (GNFP)

Fix some signature σ of relations and constants.

Syntax of GNFP[σ]

$$\varphi ::= Rt \mid Yt \mid \varphi \wedge \varphi \mid \varphi \vee \varphi \mid \exists y(\psi(\mathbf{x}y)) \mid G(\mathbf{x}) \wedge \neg\psi(\mathbf{x}) \mid \\ [\text{Ifp}_{Y,y} . G(\mathbf{y}) \wedge \varphi(\mathbf{y}, Y, \mathbf{Z})](\mathbf{t}) \quad \text{where } Y \text{ only occurs positively in } \varphi$$

where R and G are relations in σ or $=$, and \mathbf{t} is a tuple over variables and constants.

Examples

- unions of conjunctive queries (in GNF)
- frontier-guarded tgds (in GNF):
 $\forall xyz ((Rxy \wedge Ryz) \rightarrow \exists w(Twyz)) \equiv \neg\exists xyz (Rxy \wedge Ryz \wedge \neg\exists w(Twyz))$
- description logics including \mathcal{ALC} , \mathcal{ALCFIO} , \mathcal{ELJ} (in GNF)

Guarded negation fixpoint logic (GNFP)

Fix some signature σ of relations and constants.

Syntax of GNFP[σ]

$$\varphi ::= R\mathbf{t} \mid Y\mathbf{t} \mid \varphi \wedge \varphi \mid \varphi \vee \varphi \mid \exists y(\psi(\mathbf{xy})) \mid G(\mathbf{x}) \wedge \neg\psi(\mathbf{x}) \mid$$
$$[\text{Ifp}_{Y,y} \cdot G(\mathbf{y}) \wedge \varphi(\mathbf{y}, Y, \mathbf{Z})](\mathbf{t}) \quad \text{where } Y \text{ only occurs positively in } \varphi$$

where R and G are relations in σ or $=$, and \mathbf{t} is a tuple over variables and constants.

Examples

- unions of conjunctive queries (in GNF)
- frontier-guarded tgds (in GNF):
 $\forall xyz ((Rxy \wedge Ryz) \rightarrow \exists w(Twyz)) \equiv \neg\exists xyz (Rxy \wedge Ryz \wedge \neg\exists w(Twyz))$
- description logics including \mathcal{ALC} , \mathcal{ALCFIO} , \mathcal{ELI} (in GNF)
- mu-calculus, even with backwards modalities

Guarded negation fixpoint logic (GNFP)

Fix some signature σ of relations and constants.

Syntax of GNFP[σ]

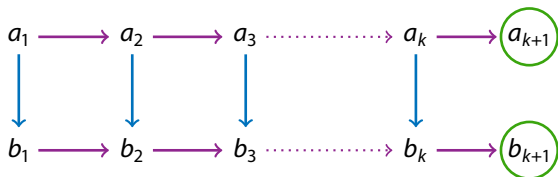
$$\varphi ::= R\mathbf{t} \mid Y\mathbf{t} \mid \varphi \wedge \varphi \mid \varphi \vee \varphi \mid \exists y(\psi(\mathbf{x}y)) \mid G(\mathbf{x}) \wedge \neg\psi(\mathbf{x}) \mid$$
$$[\text{Ifp}_{Y,y} \cdot G(\mathbf{y}) \wedge \varphi(\mathbf{y}, Y, \mathbf{Z})](\mathbf{t}) \quad \text{where } Y \text{ only occurs positively in } \varphi$$

where R and G are relations in σ or $=$, and \mathbf{t} is a tuple over variables and constants.

Examples

- unions of conjunctive queries (in GNF)
- frontier-guarded tgds (in GNF):
 $\forall xyz ((Rxy \wedge Ryz) \rightarrow \exists w(Twyz)) \equiv \neg \exists xyz (Rxy \wedge Ryz \wedge \neg \exists w(Twyz))$
- description logics including \mathcal{ALC} , \mathcal{ALCFIO} , \mathcal{ELI} (in GNF)
- mu-calculus, even with backwards modalities
- monadic Datalog

$$[\text{lfp}_{Z,xy} . Sxy \wedge \exists uv(Rxu \wedge Ryv \wedge (Zuv \vee (Pu \wedge Pv)))](xy)$$



Some nice computational properties for guarded fixpoint logics

Decidable **satisfiability** and **finite satisfiability**

(2EXPTIME in general, EXPTIME for fixed-width formulas in GFP)

[Grädel, Walukiewicz '99 ; Bárány, Segoufin, ten Cate '11; Bárány, Bojańczyk '12]

Some nice computational properties for guarded fixpoint logics

Decidable **satisfiability** and **finite satisfiability**

(2EXPTIME in general, EXPTIME for fixed-width formulas in GFP)

[Grädel, Walukiewicz '99 ; Bárány, Segoufin, ten Cate '11; Bárány, Bojańczyk '12]

Decidable **boundedness**

(given $\psi(\mathbf{y}, Y)$ positive in Y , is there $n \in \mathbb{N}$ such that for all \mathfrak{A} , $\psi_{\mathfrak{A}}^n = \psi_{\mathfrak{A}}^{n+1}$?)

[Blumensath, Otto, Weyer '14 ; Bárány, ten Cate, Otto '12 ; Benedikt, ten Cate, Colcombet, VB. '15]

Some nice computational properties for guarded fixpoint logics

Decidable **satisfiability** and **finite satisfiability**

(2EXPTIME in general, EXPTIME for fixed-width formulas in GFP)

[Grädel, Walukiewicz '99 ; Bárány, Segoufin, ten Cate '11; Bárány, Bojańczyk '12]

Decidable **boundedness**

(given $\psi(\mathbf{y}, Y)$ positive in Y , is there $n \in \mathbb{N}$ such that for all \mathfrak{A} , $\psi_{\mathfrak{A}}^n = \psi_{\mathfrak{A}}^{n+1}$?)

[Blumensath, Otto, Weyer '14 ; Bárány, ten Cate, Otto '12 ; Benedikt, ten Cate, Colcombet, VB. '15]

Constructive **interpolation** for UNFP

[Benedikt, ten Cate, VB. '15]

Why so many nice properties?

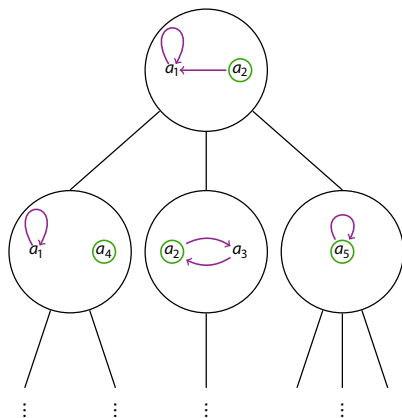
These guarded logics all have **tree-like models** (bounded tree width).

Why so many nice properties?

These guarded logics all have **tree-like models** (bounded tree width).

A structure \mathfrak{A} has **tree width** $k - 1$ if it can be covered by (overlapping) bags of size at most k , arranged in a tree t s.t.

- every fact appears in some bag in t ;
- for each element, the set of bags with this element is connected in t .



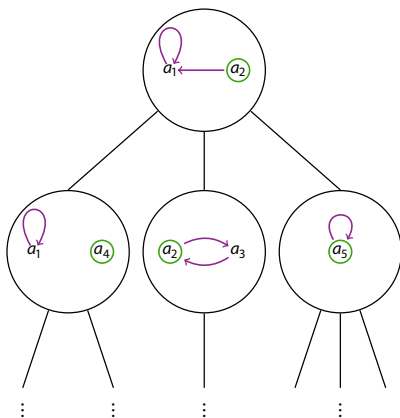
Why so many nice properties?

These guarded logics all have **tree-like models** (bounded tree width).

A structure \mathfrak{A} has **tree width** $k - 1$ if it can be covered by (overlapping) bags of size at most k , arranged in a tree t s.t.

- every fact appears in some bag in t ;
- for each element, the set of bags with this element is connected in t .

There is a natural **encoding** of these tree-like models (of some bounded tree width) as trees over a finite alphabet.



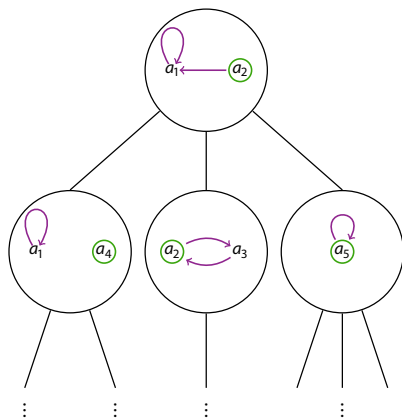
Why so many nice properties?

These guarded logics all have **tree-like models** (bounded tree width).

A structure \mathfrak{A} has **tree width** $k - 1$ if it can be covered by (overlapping) bags of size at most k , arranged in a tree t s.t.

- every fact appears in some bag in t ;
- for each element, the set of bags with this element is connected in t .

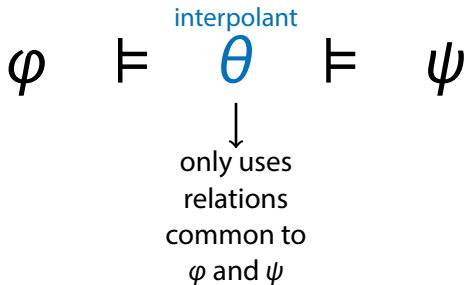
There is a natural **encoding** of these tree-like models (of some bounded tree width) as trees over a finite alphabet.



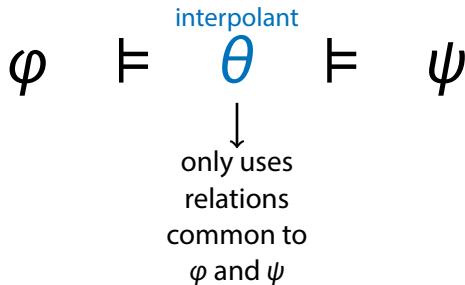
⇒ We can reason about tree encodings rather than relational structures.

$\varphi \quad \vDash \quad \psi$

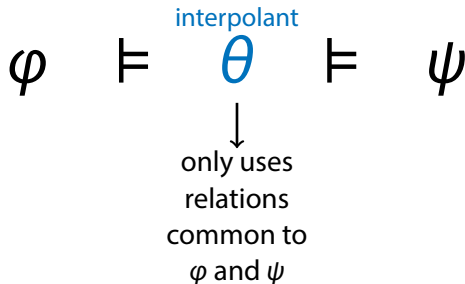
Interpolation



Interpolation



Craig interpolation: θ depends on φ and ψ



Craig interpolation: θ depends on φ and ψ

Uniform interpolation: θ depends only on φ and common signature (not on a particular ψ)

Uniform interpolation example

“ P holds at x , and from every position y where P holds, there is an R -neighbor z where P holds”

$$\begin{aligned}\varphi(x) &:= Px \wedge \forall y (Py \rightarrow \exists z (Ryz \wedge Pz)) \\ &\equiv Px \wedge \neg \exists y (Py \wedge \neg \exists z (Ryz \wedge Pz))\end{aligned}$$

Uniform interpolation example

“ P holds at x , and from every position y where P holds, there is an R -neighbor z where P holds”

$$\begin{aligned}\varphi(x) &:= Px \wedge \forall y(Py \rightarrow \exists z(Ryz \wedge Pz)) \\ &\equiv Px \wedge \neg\exists y(Py \wedge \neg\exists z(Ryz \wedge Pz))\end{aligned}$$

Uniform interpolant of φ over subsignature $\{R\}$

“there is an infinite R -path from x ”

$$\begin{aligned}&\neg[\mathbf{lfp}_{Y,y} . \forall z(Ryz \rightarrow Yz)](x) \\ \equiv &\neg[\mathbf{lfp}_{Y,y} . \neg\exists z(Ryz \wedge \neg Yz)](x)\end{aligned}$$

Why study interpolation?

- Interpolation implies several results about going from **semantic properties to syntactic properties** (e.g., Beth definability, preservation theorems, etc.)

Why study interpolation?

- Interpolation implies several results about going from **semantic properties to syntactic properties** (e.g., Beth definability, preservation theorems, etc.)
- Interpolation is related to **query rewriting** over views.

Why study interpolation?

- Interpolation implies several results about going from **semantic properties to syntactic properties** (e.g., Beth definability, preservation theorems, etc.)
- Interpolation is related to **query rewriting** over views.
- Interpolation is related to **modularity**.

Why study interpolation?

- Interpolation implies several results about going from **semantic properties to syntactic properties** (e.g., Beth definability, preservation theorems, etc.)
- Interpolation is related to **query rewriting** over views.
- Interpolation is related to **modularity**.
- Very little was known about interpolation for fixpoint logics over general relational structures, where relations can have arbitrary arity.

Theorem (D'Agostino, Hollenberg '00)

L_μ has effective uniform interpolation.

Interpolation for L_μ and UNFP

Theorem (D'Agostino, Hollenberg '00)

L_μ has effective uniform interpolation.

Let UNFP^k denote the k -variable fragment of UNFP (in normal form...).

Theorem (Benedikt, ten Cate, VB. '15)

UNFP^k has effective uniform interpolation.

UNFP has effective Craig interpolation.

Interpolation for L_μ and UNFP

Theorem (D'Agostino, Hollenberg '00)

L_μ has effective uniform interpolation.

Let UNFP^k denote the k -variable fragment of UNFP (in normal form...).

Theorem (Benedikt, ten Cate, VB. '15)

UNFP^k has effective uniform interpolation.

UNFP has effective Craig interpolation.

Proof strategy: Exploit [tree-like models](#) and ideas / results from [Grädel, Walukiewicz '99 ; Grädel, Hirsch, Otto '00 ; D'Agostino, Hollenberg '00].

Uniform interpolation for UNFP^k

Theorem (Benedikt, ten Cate, VB. '15)

UNFP^k has effective uniform interpolation.

Uniform interpolation for UNFP^k

Theorem (Benedikt, ten Cate, VB. '15)

UNFP^k has effective uniform interpolation.

Proof structure:

Relational
structures

Encodings of
tree-like models
of width k

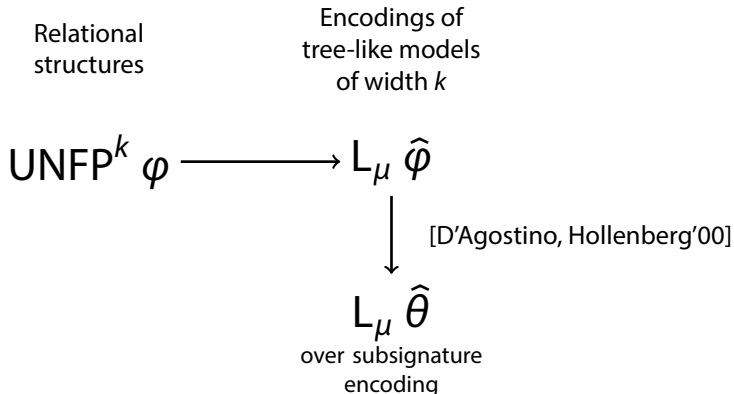
$$\text{UNFP}^k \varphi \longrightarrow L_\mu \hat{\varphi}$$

Uniform interpolation for UNFP^k

Theorem (Benedikt, ten Cate, VB. '15)

UNFP^k has effective uniform interpolation.

Proof structure:

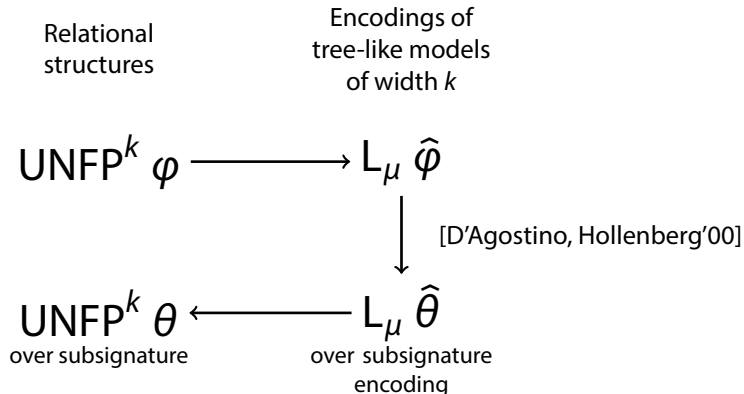


Uniform interpolation for UNFP^k

Theorem (Benedikt, ten Cate, VB. '15)

UNFP^k has effective uniform interpolation.

Proof structure:



Summary of interpolation results

UNFP has effective interpolation,
and the construction takes advantage
of its tree-like models.

	K	GF	UNF	GNF	L_μ	GFP	UNFP	GNFP
Craig interpolation	✓	✗	✓	✓	✓	✗	✓	✗

Can we go further?

Can we go further?

GNFP^{UP}: extend GNFP with **parameters in fixpoint**
(while retaining restrictions on negation).

Can we go further?

GNFP^{UP}: extend GNFP with **parameters in fixpoint**
(while retaining restrictions on negation).

“from y , it is possible to R -reach some P -element”

$$[\text{Reach-}P](y) := [\text{Ifp}_{Y,y} . \exists z (Ryz \wedge (Pz \vee Yz))](y)$$

“from y , it is possible to R -reach x ”

$$[\text{Reach-}x](y) := [\text{Ifp}_{Y,y} . \exists z (Ryz \wedge (z = x \vee Yz))](y)$$

Can we go further?

GNFP^{UP}: extend GNFP with **parameters in fixpoint**
(while retaining restrictions on negation).

“from y , it is possible to R -reach some P -element”

$$[\text{Reach-}P](y) := [\text{Ifp}_{Y,y} . \exists z (Ryz \wedge (Pz \vee Yz))](y)$$

“from y , it is possible to R -reach x ”

$$[\text{Reach-}x](y) := [\text{Ifp}_{Y,y} . \exists z (Ryz \wedge (z = x \vee Yz))](y)$$

Can we go further?

GNFP^{UP}: extend GNFP with **parameters in fixpoint**
(while retaining restrictions on negation).

“from y , it is possible to R -reach some P -element”

$$[\text{Reach-}P](y) := [\text{Ifp}_{Y,y} . \exists z (Ryz \wedge (Pz \vee Yz))](y)$$

“from y , it is possible to R -reach x ”

$$[\text{Reach-}x](y) := [\text{Ifp}_{Y,y} . \exists z (Ryz \wedge (z = x \vee Yz))](y)$$

Subsumes

C2RPQs (conjunctive 2-way regular path queries) and
MQs and GQs [Rudolph, Krötsch '13 ; Bourhis, Krötsch, Rudolph '15]

Can we go further?

GNFP^{UP}: extend GNFP with **parameters in fixpoint**
(while retaining restrictions on negation).

“from y , it is possible to R -reach some P -element”

$$[\text{Reach-}P](y) := [\text{Ifp}_{Y,y} . \exists z (Ryz \wedge (Pz \vee Yz))](y)$$

“from y , it is possible to R -reach x ”

$$[\text{Reach-}x](y) := [\text{Ifp}_{Y,y} . \exists z (Ryz \wedge (z = x \vee Yz))](y)$$

Subsumes

C2RPQs (conjunctive 2-way regular path queries) and
MQs and GQs [Rudolph, Krötsch '13 ; Bourhis, Krötsch, Rudolph '15]

But still has **tree-like models!**

GNFP^{UP}: extend GNFP with parameters in fixpoint
(while retaining restrictions on negation).

GNFP^{UP}: extend GNFP with parameters in fixpoint
(while retaining restrictions on negation).

Theorem (Benedikt, Bourhis, VB. unpublished)

Satisfiability is decidable for $\varphi \in \text{GNFP}^{\text{UP}}$ in $(n + 2)$ -EXPTIME,
where n is “nesting depth of UCQ-shaped formulas with parameters” in φ .

Boundedness is decidable for $\varphi \in \text{GNFP}^{\text{UP}}$.

GNFP^{UP}: extend GNFP with parameters in fixpoint
(while retaining restrictions on negation).

Theorem (Benedikt, Bourhis, VB. unpublished)

Satisfiability is decidable for $\varphi \in \text{GNFP}^{\text{UP}}$ in $(n + 2)$ -EXPTIME,
where n is “nesting depth of UCQ-shaped formulas with parameters” in φ .

Boundedness is decidable for $\varphi \in \text{GNFP}^{\text{UP}}$.

Open questions

Does GNFP^{UP} have interpolation?

Is finite satisfiability decidable for GNFP^{UP}?

Guarded fixpoint logics are expressive logics
with nice computational properties
coming from their tree-like models.

Examples

Expressible in GNF

R is symmetric

$$\begin{aligned} & \forall xy(Rxy \rightarrow Ryx) \\ \equiv & \neg \exists xy(Rxy \wedge \neg Ryx) \end{aligned}$$

Every element has an R -successor

$$\begin{aligned} & \forall x(\exists y(Rxy)) \\ \equiv & \neg \exists x(\neg \exists y(Rxy)) \end{aligned}$$

Every element is on R -cycle of length 3

$$\begin{aligned} & \forall x \exists yz(Rxy \wedge Ryz \wedge Rzx) \\ \equiv & \neg \exists x(\neg \exists yz(Rxy \wedge Ryz \wedge Rzx)) \end{aligned}$$

Not expressible in GNF

R is total

$$\begin{aligned} & \forall xy(Rxy \vee Ryx) \\ \equiv & \neg \exists xy(\neg Rxy \wedge \neg Ryx) \end{aligned}$$

Every element has a unique R -successor

$$\begin{aligned} & \forall x \exists y(Rxy \wedge \forall z(Rxz \rightarrow y = z)) \\ \equiv & \neg \exists x(\neg \exists y(Rxy \wedge \neg \exists z(Rxz \wedge y \neq z))) \end{aligned}$$

Every element is on R -cycle

$$\begin{aligned} & \forall x [\text{lfp}_{Y,y}^x. \exists z(Ryz \wedge (z = x \vee Yz))](x) \\ \equiv & \neg \exists x [\text{lfp}_{Y,y}^x. \exists z(Ryz \wedge (z = x \vee Yz))](x) \end{aligned}$$

Examples, continued

Expressible in GNFP

R is well-founded:

$$\begin{aligned} & \forall yz (Ryz \rightarrow [\text{lfp}_{Y,y} \cdot \forall x (Rxy \rightarrow Yx)](y)) \\ \equiv & \neg \exists yz (Ryz \wedge \neg [\text{lfp}_{Y,y} \cdot \neg \exists x (Rxy \wedge \neg Yx)](y)) \end{aligned}$$