

Bounded Treewidth and the Infinite Core Chase

Complications and Workarounds toward Decidable Querying

Jean-François Baget¹ Marie-Laure Mugnier¹ Sebastian Rudolph²

¹*LIRMM, Inria, University of Montpellier, CNRS, France*

²*Computational Logic Group, TU Dresden*

2023 ACM SIGMOD/PODS @Seattle, WA, USA



Query Answering with Existential Rules

Basic Problem (Entailment)

INPUT: a knowledge base $K = (F, \Sigma)$ where F is a database instance and Σ is a finite set of existential rules, a (Boolean) conjunctive query Q .

QUESTION: does \mathcal{K} entail Q , i.e. $F, \Sigma \models Q$?

- Atomset:** a possibly infinite (countable) set of atoms over constants and variables (no equality, no function symbol);
- Database (Instance):** a finite atomset;
- (Existential) Rule (or tgd):** a pair of finite atomsets whose general logic form is $\forall \vec{X} \forall \vec{Y} (body(\vec{X}, \vec{Y}) \rightarrow \exists \vec{Z} head(\vec{Y}, \vec{Z}))$;
- (Boolean Conjunctive) Query:** a finite atomset.

Object	Example	Logical form
Database	$p(a, b), q(b, c)$	$p(a, b) \wedge q(b, c)$
Rule	$p(X, Y), q(Y, Z) \rightarrow r(X, T, Z), s(T)$	$\forall X \forall Y \forall Z (p(X, Y) \wedge q(Y, Z) \rightarrow \exists T (r(X, T, Z) \wedge s(T)))$
Query	$r(a, U, b), p(a, V), q(W, b)$	$\exists U \exists V \exists W (r(a, U, b) \wedge p(a, V) \wedge q(W, b))$

The Great AIJ Blunder of 2011 (BagLecMugSal11)

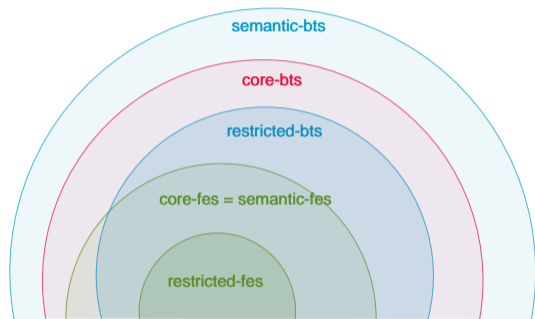


Figure: A cartography of some (abstract) decidable subclasses for the entailment problem, following (BagLecMugSal11)

- 1 **Preliminaries:** why is the proof of decidability for *core-bts* wrong ?
- 2 **The steepening staircase:** the class *core-bts* is misplaced !
- 3 **Robust aggregations:** a new proof of decidability for the *core-bts* class.
 - a novel way to define the result of an infinite chase;
 - the need to consider *finitely-universal* models instead of the usual *universal models*.

Preliminary Notions

Semantic Definition of Finite Expansion Sets (fes)

- 1 **Trigger:** a (K -)trigger for an atomset F is a pair $t = (R, \pi)$ where $R \in \Sigma$ and π maps $body(R)$ into F . It is *satisfied* when π extends to map $body(R) \cup head(R)$ into F .
- 2 **Model:** an atomset I (seen as an interpretation) is a *model* (of K) when it is a model of F and all K -triggers for I are satisfied.
- 3 **Universality:** an atomset is *universal* (for K) when it maps to every model of K .
- 4 **(BCQ) Representative:** a (BCQ)-representative of K is an atomset I such that, for any Q , we have $K \models Q \Leftrightarrow I \models Q$.

Theorem (Universal Models)

If an atomset is a universal model of K , then it is a BCQ representative of K .

- 5 **Semantic fes:** a set of rules Σ belongs to the (decidable but unrecognizable) *semantic fes* class when, for every F , (F, Σ) admits a finite universal model.

Semantic Definition of Bounded Treewidth Sets (bts)

Theorem (Treewidth and decidability, (Cou90) + (BagLecMugSal11))

Entailment is decidable for KBs admitting a universal model of finite treewidth.

- 2 **Treewidth:** the *treewidth* of an atomset F measures its similarity to a tree. If F is a tree, then $tw(F) = 1$. If F contains a grid of unbounded size, then $tw(F) = +\infty$.
- 3 **Semantic bts:** a set of rules Σ belongs to the (decidable) *semantic bts* class when, for every F , (F, Σ) admits an universal model of finite treewidth.
- 4 semantic-fes \subset semantic-bts

Theorem (Compactness of treewidth, Thomas88thetree-width)

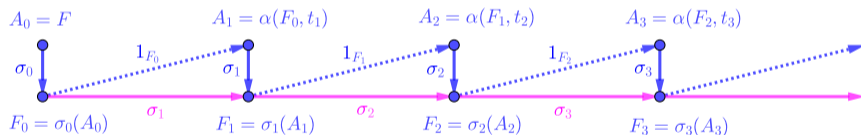
If every finite subset B of an atomset A has treewidth $tw(B) \leq k$, then $tw(A) \leq k$.

Derivations and their Results (1)

- ① **Rule application:** let $t = (R, \pi)$ be a trigger in F . Then the *application* of t on F produces the atomset $\alpha(F, t) = F \cup \pi^{safe}(\text{head}(R))$.

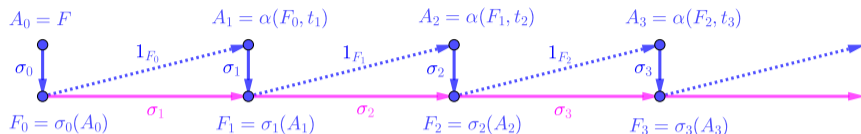
Proposition (Properties of Rule Application)

A trigger t for F is satisfied in $\alpha(F, t)$. Moreover, if F is universal, then $\alpha(F, t)$ is universal.



- ② **Derivation:** a possibly infinite sequence $\mathcal{D} = (F_i)$ where $F_0 = \sigma_0(F)$ and $F_i = \sigma_i(\alpha(F_{i-1}, t_i))$, the σ_i being endomorphisms.
- ③ **Fairness:** \mathcal{D} is said *fair* when, for any trigger (R, π) for some F_i , there is some F_j in which $(R, \sigma_j^i \circ \pi)$ is satisfied.

Derivations and their Results (2)



- 1 Finite result:** if \mathcal{D} is a finite derivation, then its *finite result* \mathcal{D}^+ is its last atomset.
- 2 Natural aggregation:** the *natural aggregation* of a (possibly infinite) derivation $\mathcal{D} = (F_i)_{i \in \mathcal{J}}$ is the (possibly infinite) atomset $\mathcal{D}^* = \cup_{i \in \mathcal{J}} F_i$.

Theorem (Finite result)

The finite result \mathcal{D}^+ of a fair derivation is a finite universal model.

Theorem (Natural aggregation)

The natural aggregation \mathcal{D}^ of a fair monotonic derivation \mathcal{D} is a (possibly infinite) universal model. If \mathcal{D} is non-monotonic, then \mathcal{D}^* is a universal BCQ representative, but not necessarily a model.*

The Restricted and Core Chases: the Terminating Case

- 1 **Restricted chase (fkmp05):** the *restricted chase* is a fair derivation that only applies unsatisfied triggers and whose endomorphisms σ_i are the identity. The restricted chase is *monotonic*.
- 2 **Core:** a *finite atomset* is a *core* when there is no homomorphism into one of its strict subsets. Every finite atomset maps to a subset which is a core.
- 3 **Core chase (DBLP:conf/pods/DeutschNR08):** the *core chase* is a fair derivation that only applies unsatisfied triggers and whose endomorphisms σ_i map A_i to a core. The core chase is not always monotonic.
- 4 **Restricted-fes:** a set of rules Σ belongs to the (decidable) *restricted-fes* class when, for every F , the restricted chase halts on (F, Σ) .
- 5 **Core-fes:** a set of rules Σ belongs to the (decidable) *core-fes* class when, for every F , the core chase halts on (F, Σ) .
- 6 $\text{restricted-fes} \subset \text{core-fes} = \text{semantic-fes}$

The Restricted and Core Chases: the Bounded Treewidth Case

- ① **Bounded treewidth:** a derivation \mathcal{D} has *bounded treewidth* k when $\forall F_i, tw(F_i) \leq k$.

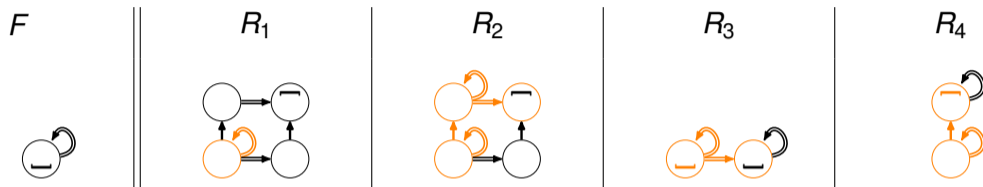
Theorem (Treewidth and monotonic derivations)

If \mathcal{D} is a monotonic derivation with bounded treewidth k , then $tw(\mathcal{D}^*) \leq k$.

- | | |
|---|--|
| <p>① the <i>natural aggregation</i> \mathcal{D}^* of a restricted chase is a universal model.</p> <p>② the <i>natural aggregation</i> \mathcal{D}^* of a restricted chase of bounded treewidth is an atomset of finite treewidth.</p> <p>③ Restricted-bts: a set of rules Σ belongs to the (decidable) <i>restricted-bts</i> class when, for every F, the restricted chase from (F, Σ) has bounded treewidth.</p> | <p>① the <i>natural aggregation</i> \mathcal{D}^* of a core chase is not necessarily a model.</p> <p>② the <i>natural aggregation</i> \mathcal{D}^* of a core chase of bounded treewidth may not have finite treewidth.</p> <p>③ No reason for core-bts decidability: a set of rules Σ belongs to the <i>core-bts</i> class when, for every F, the core chase from (F, Σ) has bounded treewidth.</p> |
|---|--|

The Steepening Staircase

The Steepening Staircase: Presentation of the KB

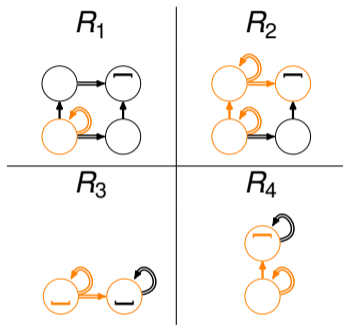


1 **Facts:** $d(X), h(X, X)$.

2 **Rules:**

- 1 $h(X, X) \rightarrow \exists Y \exists Z \exists T (v(X, Y), h(Y, Z), u(Z), v(T, Z), h(X, T))$.
- 2 $h(X, X), v(X, Y), h(Y, Y), h(Y, Z) \rightarrow \exists T h(X, T), v(T, Z), u(Z)$.
- 3 $d(X), h(X, X), h(X, Y) \rightarrow d(Y), h(Y, Y)$.
- 4 $h(X, X), v(X, Y), u(Y) \rightarrow h(Y, Y)$.

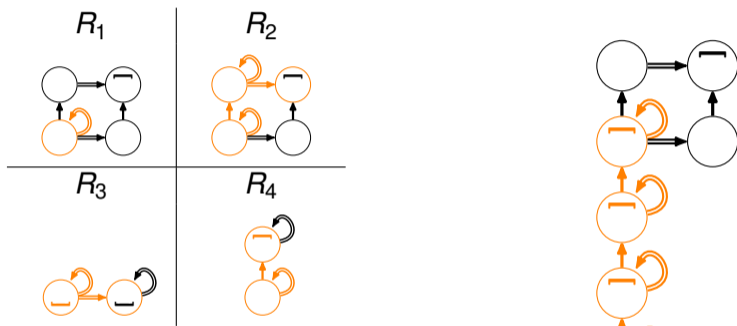
The Steepening Staircase: Elementary Step of a Derivation



- Elementary step of the derivation:**
from a C_k we build a S_k , containing a C_{k+1}

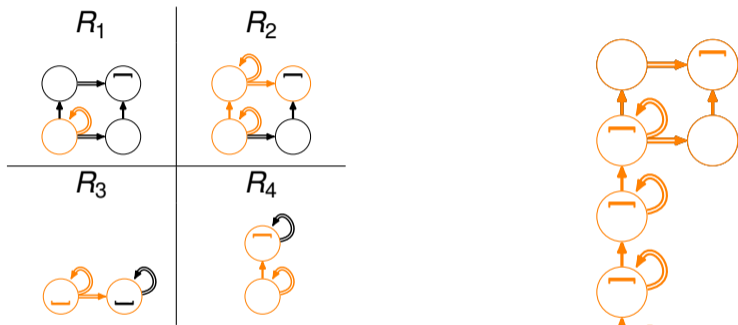


The Steepening Staircase: Elementary Step of a Derivation



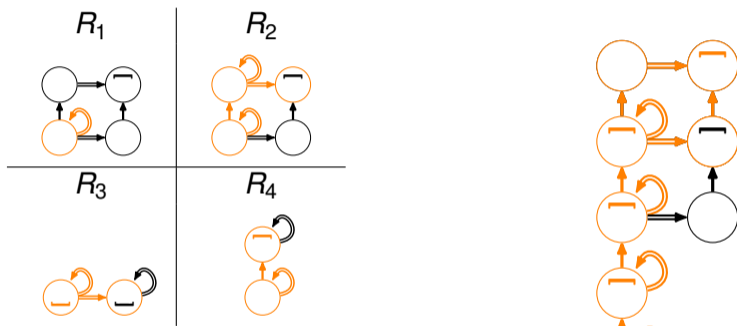
- Elementary step of the derivation:**
from a C_k we build a S_k , containing a C_{k+1}

The Steepening Staircase: Elementary Step of a Derivation



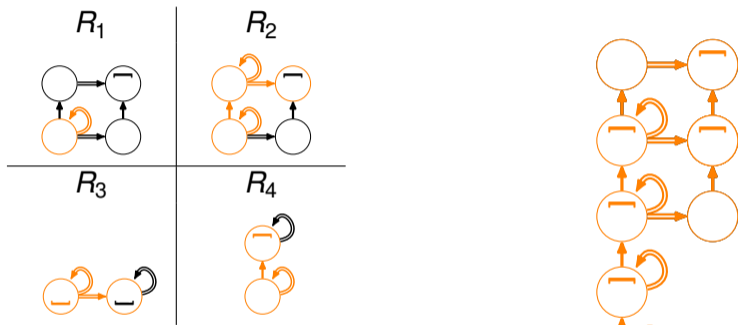
- Elementary step of the derivation:**
from a C_k we build a S_k , containing a C_{k+1}

The Steepening Staircase: Elementary Step of a Derivation



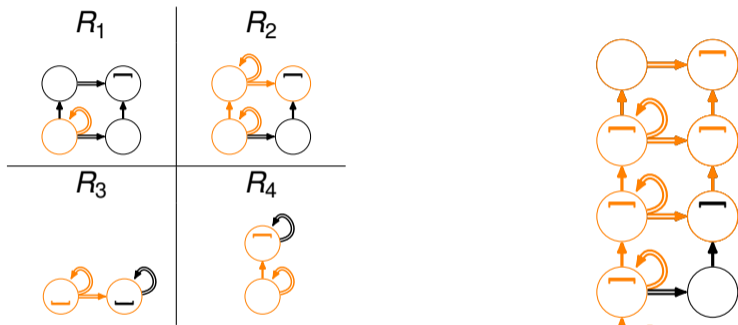
- Elementary step of the derivation:**
from a C_k we build a S_k , containing a C_{k+1}

The Steepening Staircase: Elementary Step of a Derivation



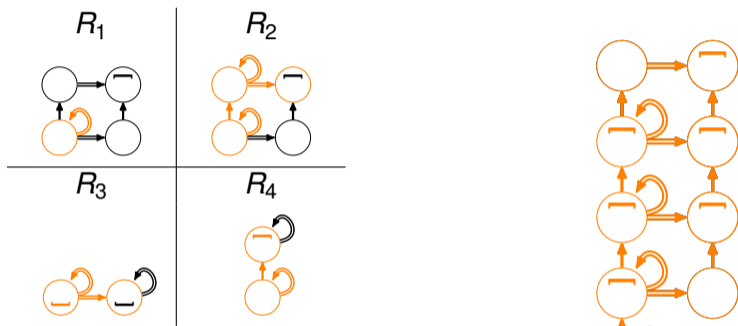
- Elementary step of the derivation:**
from a C_k we build a S_k , containing a C_{k+1}

The Steepening Staircase: Elementary Step of a Derivation



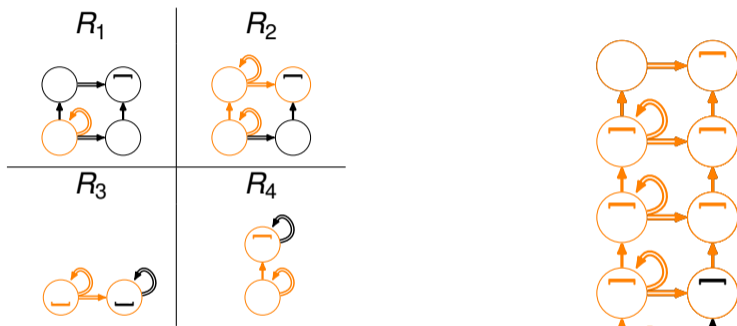
- Elementary step of the derivation:**
from a C_k we build a S_k , containing a C_{k+1}

The Steepening Staircase: Elementary Step of a Derivation



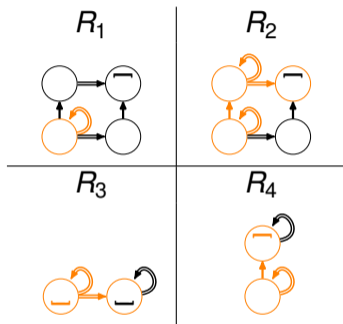
- Elementary step of the derivation:**
from a C_k we build a S_k , containing a C_{k+1}

The Steepening Staircase: Elementary Step of a Derivation

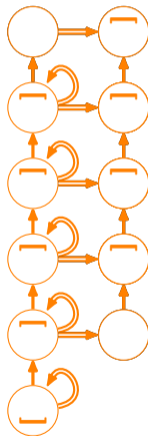


- Elementary step of the derivation:** from a C_k we build a S_k , containing a C_{k+1}

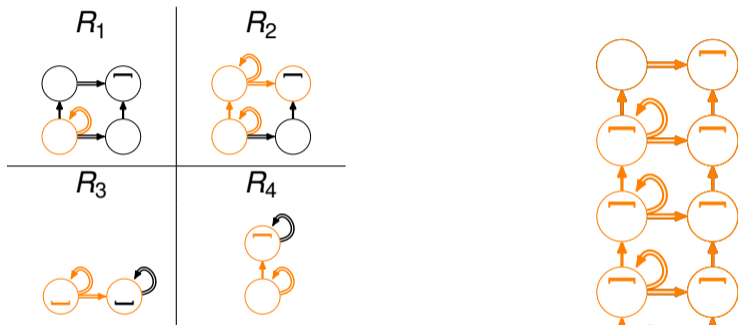
The Steepening Staircase: Elementary Step of a Derivation



- Elementary step of the derivation:**
from a C_k we build a S_k , containing a C_{k+1}

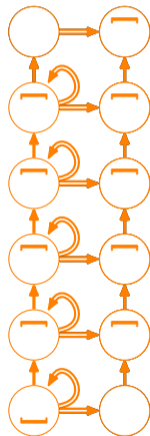
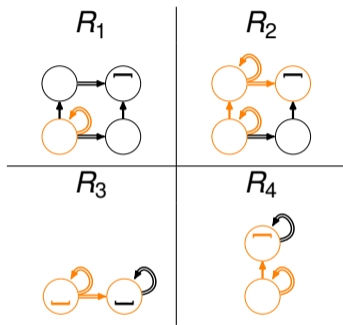


The Steepening Staircase: Elementary Step of a Derivation



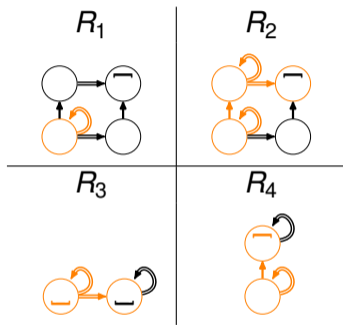
- Elementary step of the derivation:**
from a C_k we build a S_k , containing a C_{k+1}

The Steepening Staircase: Elementary Step of a Derivation

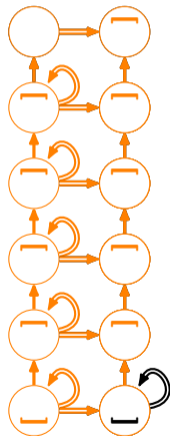


- Elementary step of the derivation:**
from a C_k we build a S_k , containing a C_{k+1}

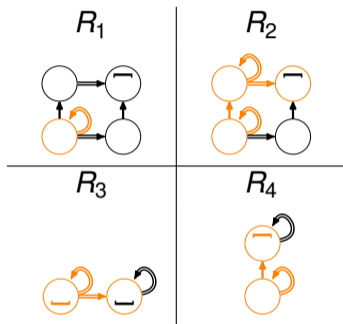
The Steepening Staircase: Elementary Step of a Derivation



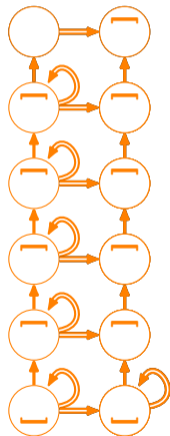
- Elementary step of the derivation:**
from a C_k we build a S_k , containing a C_{k+1}



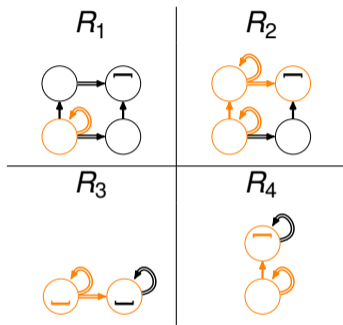
The Steepening Staircase: Elementary Step of a Derivation



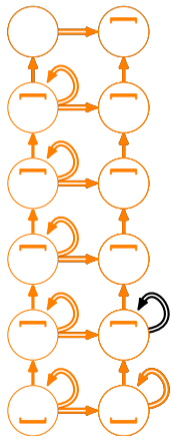
- Elementary step of the derivation:**
from a C_k we build a S_k , containing a C_{k+1}



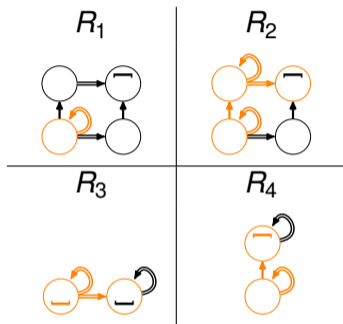
The Steepening Staircase: Elementary Step of a Derivation



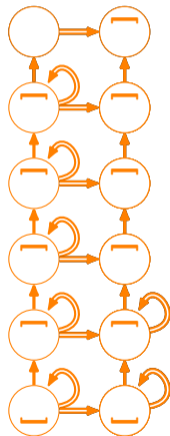
- 1 **Elementary step of the derivation:**
from a C_k we build a S_k , containing a
 C_{k+1}



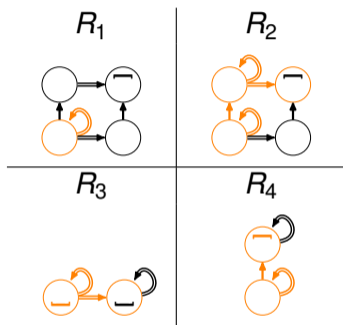
The Steepening Staircase: Elementary Step of a Derivation



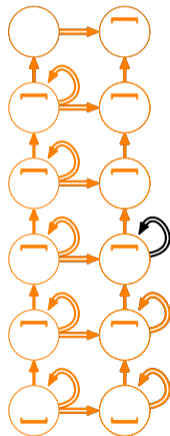
- Elementary step of the derivation:**
from a C_k we build a S_k , containing a C_{k+1}



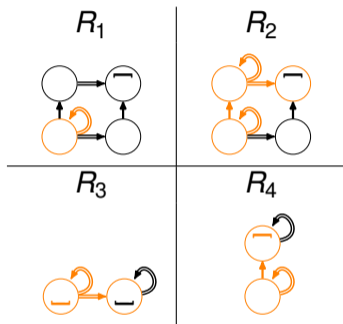
The Steepening Staircase: Elementary Step of a Derivation



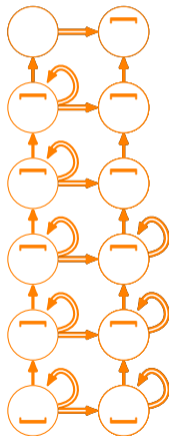
- Elementary step of the derivation:**
from a C_k we build a S_k , containing a C_{k+1}



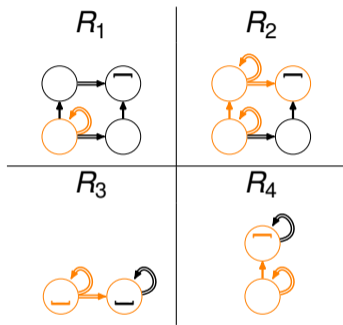
The Steepening Staircase: Elementary Step of a Derivation



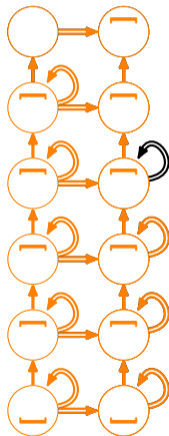
- Elementary step of the derivation:**
from a C_k we build a S_k , containing a C_{k+1}



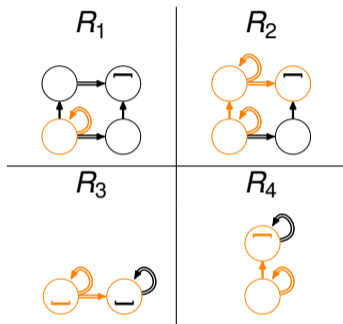
The Steepening Staircase: Elementary Step of a Derivation



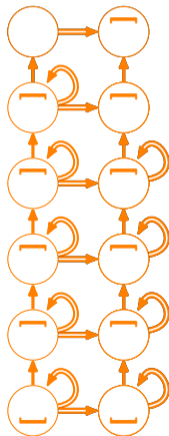
- Elementary step of the derivation:**
from a C_k we build a S_k , containing a C_{k+1}



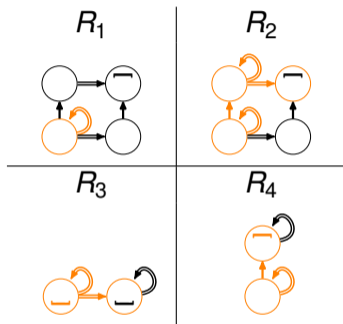
The Steepening Staircase: Elementary Step of a Derivation



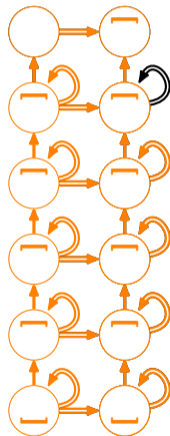
- 1 **Elementary step of the derivation:**
from a C_k we build a S_k , containing a
 C_{k+1}



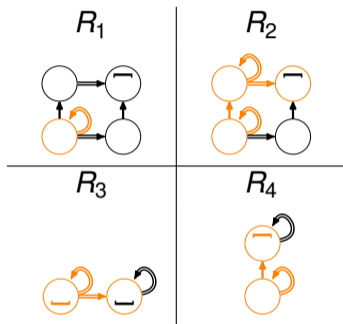
The Steepening Staircase: Elementary Step of a Derivation



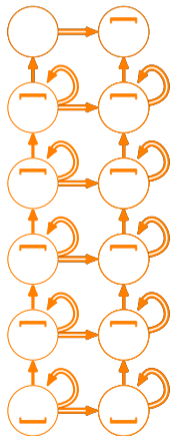
- Elementary step of the derivation:**
from a C_k we build a S_k , containing a C_{k+1}



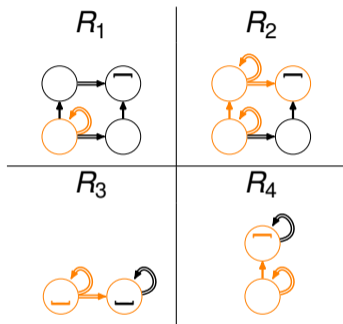
The Steepening Staircase: Elementary Step of a Derivation



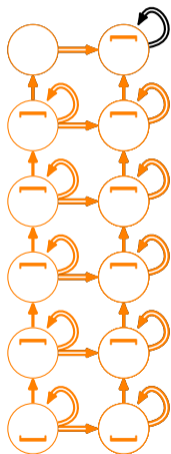
- 1 **Elementary step of the derivation:**
from a C_k we build a S_k , containing a
 C_{k+1}



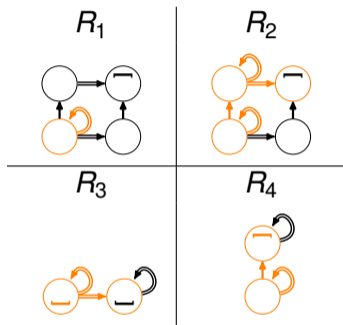
The Steepening Staircase: Elementary Step of a Derivation



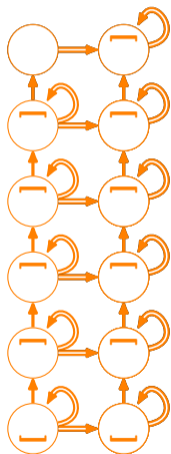
- Elementary step of the derivation:**
from a C_k we build a S_k , containing a C_{k+1}



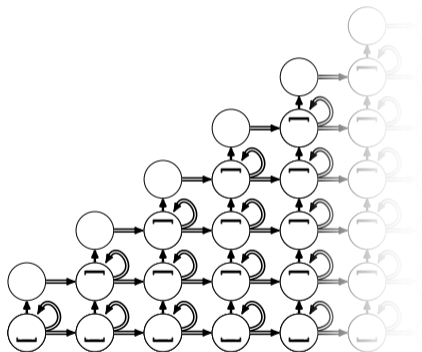
The Steepening Staircase: Elementary Step of a Derivation



- Elementary step of the derivation:**
from a C_k we build a S_k , containing a C_{k+1}

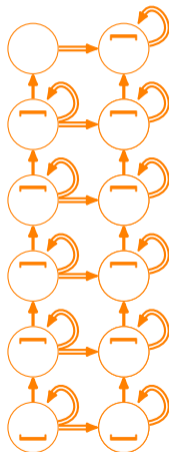


The Steepening Staircase: Universal Models of Infinite Treewidth



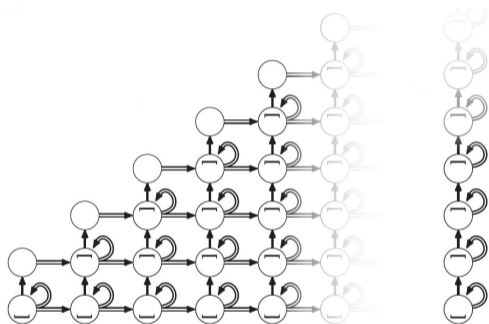
- 1 **Restricted chase:** the natural aggregation of a restricted chase \mathcal{D}^* is a universal model of infinite treewidth (grids of unbounded size).
- 2 **Moreover,** every universal model of the steepening staircase KB has infinite treewidth (see paper).

The Steepening Staircase: a Core Chase of Bounded Treewidth



- 1 **Core of a S_k :** the core of a S_k is a C_{k+1} . All atomsets built from a C_k to S_k are cores.
- 2 **Core chase:** the atomsets along the core chase have treewidth between 1 and 2.
- 3 **Treewidth:** the natural aggregation of the core chase is the same as the one obtained from the restricted chase, and has thus infinite treewidth.
- 4 **Consequence:** core-bts $\not\subseteq$ semantic-bts

Finitely-Universal Models



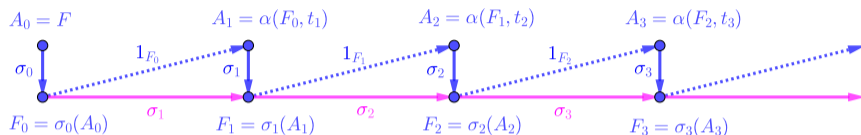
- 1 the infinite column C_∞ is a model of the steepening staircase KB, but it is not universal.
- 2 **Finite-universality:** an atomset A is *finitely-universal* when every subset of A is universal.
- 3 C_∞ is finitely-universal.

Theorem ((BCQ) representative)

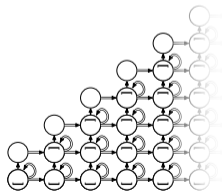
A finitely-universal model of a KB \mathcal{K} is a (BCQ) representative of \mathcal{K} .

Robust Aggregations

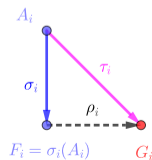
Robust Renaming



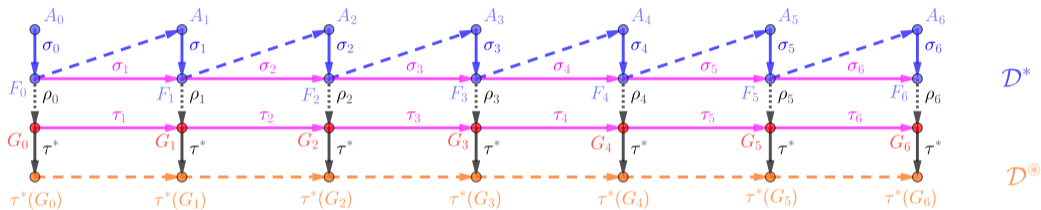
- 1 We want to define the **result** as $\cup_{i \in \mathbb{N}} \sigma^*(F_i)$, but it doesn't work !



- 1 **Robust renaming:** if $X \in \text{vars}(F_i)$, we define $\rho_i(X) = \min(\sigma_i^{-1}(X))$
- 2 ρ_i is an **isomorphism** and $\tau_i = \rho_i \circ \sigma_i$ is such that, for any $X \in \text{vars}(A_i)$, $\tau_i(X) \leq X$.



Robust Aggregation



- 1 Apply the robust renaming all along the derivation. See that $\tau_i = \rho_i \circ \sigma_i$ is also a homomorphism from G_{i-1} to G_i , and that G_i is isomorphic to F_i .
- 2 **The τ_i are finitely morphing:** if X is a variable in F_i , there is $j \geq i$ such that for any $r \geq j$, $\tau_i^j(X) = \tau_i^r(X)$. We can thus **define** $\tau^*(X) = \tau_i^j(X)$.
- 3 **Robust aggregation:** if \mathcal{D} is a derivation, we call $\mathcal{D}^* = \cup_{i \in \mathcal{J}} \tau^*(G_i)$ its *robust aggregation*.

Main Properties of Robust Aggregation

Theorem (Model)

\mathcal{D}^* is a model.

- 1 \mathcal{D}^* is not a model for nonmonotonic derivations.

Theorem (Finite-universality)

\mathcal{D}^* is finitely-universal.

- 2 \mathcal{D}^* is always universal.

Theorem

If \mathcal{D} is a derivation with bounded treewidth k , then \mathcal{D}^* has treewidth $\leq k$.

- 3 \mathcal{D}^* may have infinite treewidth (see steepening staircase).

Finishing touches

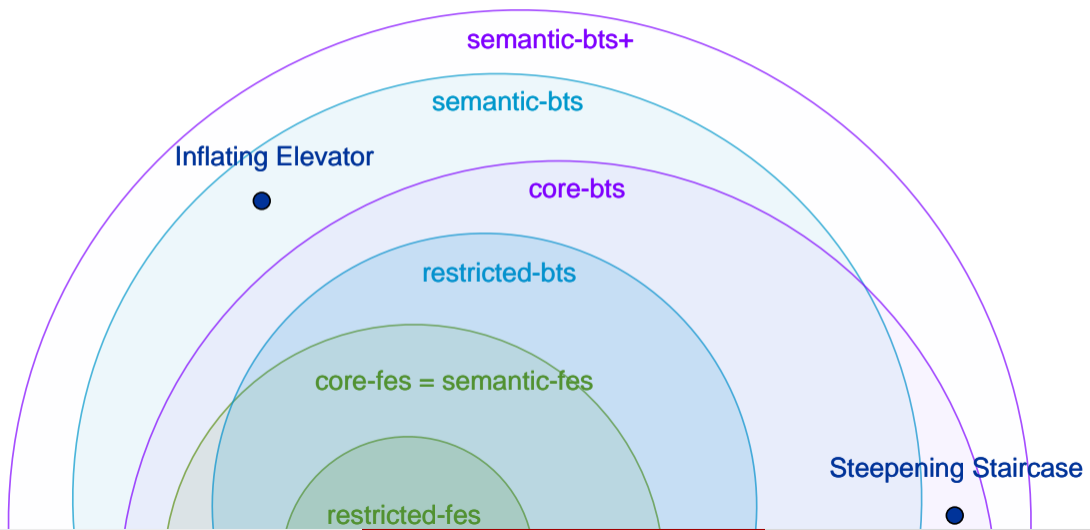
- 1 If \mathcal{D} is a chase with bounded treewidth k , then \mathcal{D}^{\otimes} is a finitely-universal model (and thus a BCQ representative) of treewidth $\leq k$.

Theorem (Basically (**Cou90**) + (**BagLecMugSal11**) + (this paper))

CQ entailment is decidable for KBs admitting a finitely-universal model of finite treewidth.

- 2 **Semantic-bts+**: a set of rules Σ belongs to the (decidable) *semantic-bts+* class when, for every F , (F, Σ) admits a finitely-universal model of finite treewidth.
- 3 this is now a proof of decidability of **core-bts** !

The New Map of Abstract Decidable Classes

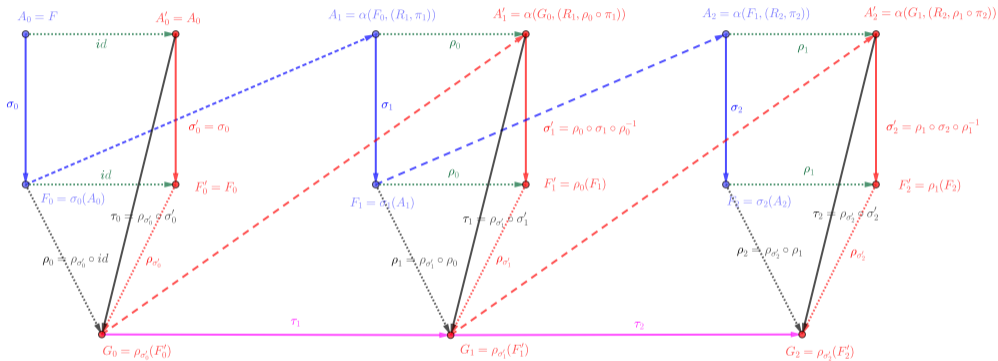


References I

Part I

Appendix

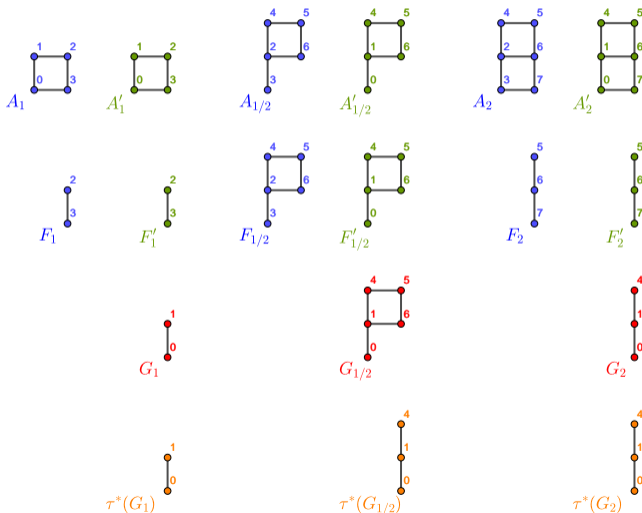
How to reach that goal (3): wrapping it up



Theorem (Finitely morphing, monotonicity)

The τ_i are finitely morphing, allowing to define τ^* . The $\tau^*(G_i)$ are **monotonic**.

The staircase example: a closer look



Main properties of robust aggregation

Theorem (Model)

\mathcal{D}^* is a model.

- 1 \mathcal{D}^* is not a model for nonmonotonic derivations.

Theorem (Finite universality)

\mathcal{D}^* is finitely universal.

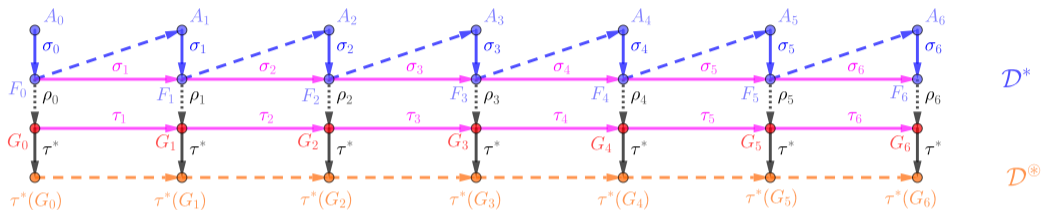
- 2 \mathcal{D}^* is always universal.

Theorem ((BCQ) representative)

A finitely universal model of a KB \mathcal{K} is a (BCQ) representative of \mathcal{K} .

- 3 The natural aggregation \mathcal{D}^* is also a (BCQ) representative.

Preliminary: a kind of monotonicity

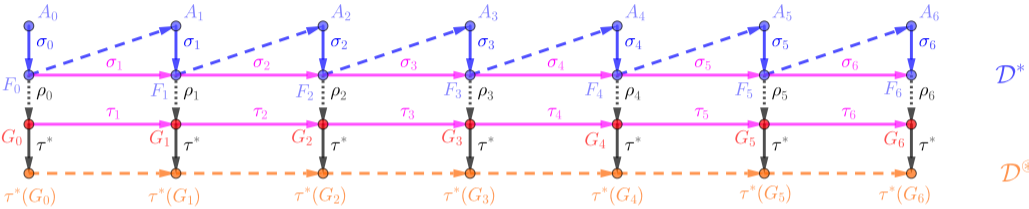


- 1 **Where is Wally ?** if $(B_i)_{i \in \mathcal{I}}$ is a **monotonic** sequence of atomsets, then for any **finite** subset A of $\cup_{i \in \mathcal{I}} B_i$, there exists $j \in \mathcal{I}$ such that, for any $k \geq j \in \mathcal{I}$, $A \subseteq B_k$.
- 2 **Monotonic derivations:** if $\mathcal{D} = (F_i)_{i \in \mathcal{I}}$ is a monotonic derivation and A is a finite subset of \mathcal{D}^* , then there exists $j \in \mathcal{I}$ such that, for any $k \geq j \in \mathcal{I}$, $A \subseteq F_k$.

Lemma (Where is Wally)

If A is a finite subset of \mathcal{D}^* , then there exists $j \in \mathcal{I}$ such that, for any $k \geq j \in \mathcal{I}$, $A \subseteq G_k$.

\mathcal{D}^{\otimes} is finitely universal



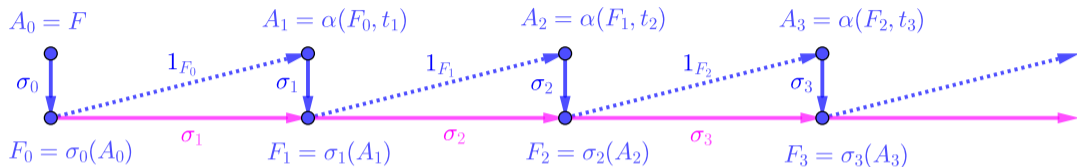
Theorem (Finite universality)

\mathcal{D}^{\otimes} is finitely universal.

- 1 if A is a finite subset of \mathcal{D}^{\otimes} , then there exists G_i such that $A \subseteq G_i$ (Wally lemma).
- 2 since G_i is isomorphic to F_i and F_i universal, then G_i is universal.
- 3 as a subset of a universal atomset, A is universal



Preliminary: treewidth and monotonic derivations

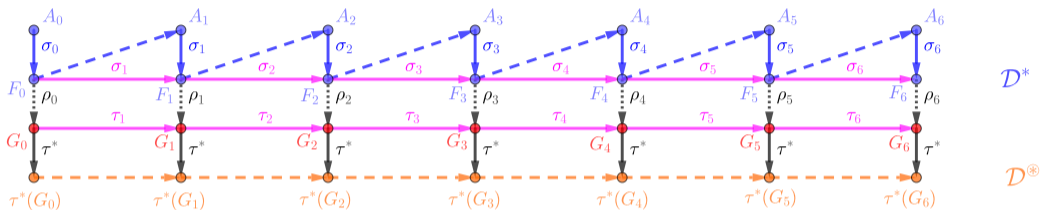


Theorem

If \mathcal{D} is a monotonic derivation with bounded treewidth k , then \mathcal{D}^* has treewidth $\leq k$.

- 1 Let us consider any finite subset A of \mathcal{D}^* . Since \mathcal{D} is **monotonic**, there exists $i \in \mathcal{I}$ such that $A \subseteq F_i$ (where is Wally).
- 2 Since $A \subseteq F_i$, we have $tw(A) \leq tw(F_i)$, and since \mathcal{D} has bounded treewidth k , we have $tw(F_i) \leq k$. Thus $tw(A) \leq k$.
- 3 We conclude with the compactness theorem: $tw(\mathcal{D}^*) \leq k$ □

Treewidth and robust aggregation



Theorem

If \mathcal{D} is a derivation with bounded treewidth k , then \mathcal{D}^* has treewidth $\leq k$.

- 1 Let us consider any finite subset A of \mathcal{D}^* . There is $i \in \mathbb{N}$ st $A \subseteq G_i$ (Wally lemma).
- 2 Since $A \subseteq G_i$, we have $tw(A) \leq tw(G_i)$, since G_i is isomorphic to F_i we have $tw(G_i) = tw(F_i)$, and since \mathcal{D} has btw k , we have $tw(F_i) \leq k$. Thus $tw(A) \leq k$.
- 3 We conclude with the compactness theorem: $tw(\mathcal{D}^*) \leq k$ □

Finishing touches (1)

- 1 If \mathcal{D} is a derivation with intermittent bounded treewidth k , then \mathcal{D}^{\otimes} is a finitely universal model (and thus a BCQ representative) of treewidth $\leq k$.

Theorem (Basically (Cou90) + (BagLecMugSal11))

CQ entailment is decidable for knowledge bases admitting a finitely universal model of finite treewidth.

- 2 A ruleset \mathcal{R} is said **fes** when, for every F , (F, \mathcal{R}) admits a finite universal model. It is said **bts** when, for every F , there is a **monotonic** derivation from (F, \mathcal{R}) (,e.g. a *restricted chase*) having uniformly bounded treewidth.
- 3 CQ entailment is decidable for KBs having fes or bts rulesets.
- 4 fes and bts are not comparable.

Quasimodels

- 1 \mathcal{D}^* is a finitely universal model.
- 2 a finitely universal model is a (BCQ) representative
- 3 \mathcal{D}^* is a universal (not a model), but a (BCQ) representative.

Objective

Find a nice characterization of a **quasimodel** such that:

- \mathcal{D}^* is a quasimodel
 - a universal quasimodel is a (BCQ) representative
-
- 1 is a finitely universal quasimodel a (BCQ) representative ?
 - 2 is CQ entailment decidable when \mathcal{K} admits a finitely universal quasimodel of finite treewidth?
 - 3 are all (BCQ) representatives finitely universal quasimodels?

Semantic BTS

- ① A ruleset \mathcal{R} is said **cci-bts** when, for every F , there is a derivation from (F, \mathcal{R}) (*e.g.* a *core chase*) having intermittent bounded treewidth.
- ② A ruleset \mathcal{R} is said **sem-bts** when, for every F , there exists a finitely universal model of (F, \mathcal{R}) with finite treewidth.
- ③ the **magic staircase** rules are sem-bts, but not cci-bts.

Remark

In the magic staircase core derivation, neither the F_i nor the G_i have (uniform or intermittent) bounded treewidth. However, the $\tau^(G_i)$ have bounded treewidth.*

- ① see that if the $\tau^*(G_i)$ have intermittent bounded treewidth, then \mathcal{D}^* has finite treewidth; this leads to a new decidable class $\text{cci-bts} \subset \mathbf{wf\text{-}cci\text{-}bts} \subseteq \text{sem-bts}$.
- ② wouldn't it be nice to have $\text{wf-cci-bts} = \text{sem-bts}$?

