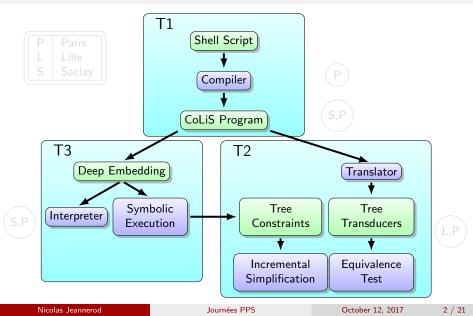
Unix filesystem and graph constraints

Nicolas Jeannerod

INSTITUT DE RECHERCHE EN INFORMATIQUE FONDAMENTALE

Journées PPS, October 12, 2017

The CoLiS project



The CoLiS project

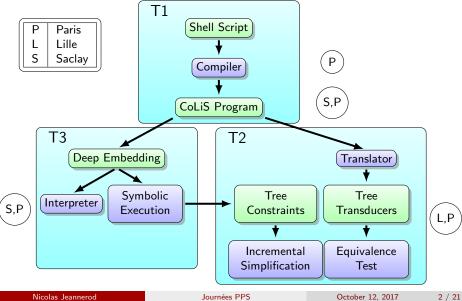
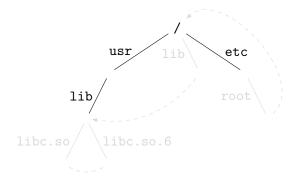


Table of Contents

1. Description of file systems

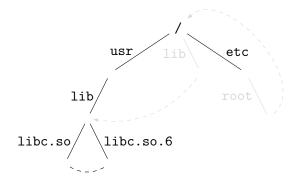
- Unix file system
- Static description
- Directory update

- 2. Tree constraints
 - Definitions
 - Basic constraints
 - Existential and first order constraints

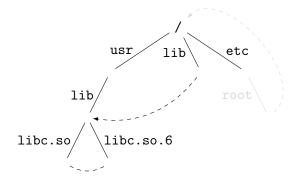


Basically a tree with labelled nodes and edges;

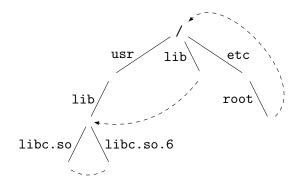
- There can be sharing at the leafs (hard link between files);
- There can be pointers to other parts of the tree (symbolic links) which may form cycles.



- Basically a tree with labelled nodes and edges;
- There can be sharing at the leafs (hard link between files);
- There can be pointers to other parts of the tree (symbolic links) which may form cycles.



- Basically a tree with labelled nodes and edges;
- There can be sharing at the leafs (hard link between files);
- There can be pointers to other parts of the tree (symbolic links) which may form cycles.



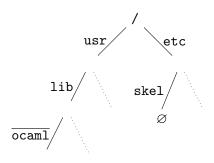
- Basically a tree with labelled nodes and edges;
- There can be sharing at the leafs (hard link between files);
- There can be pointers to other parts of the tree (symbolic links) which may form cycles.

Table of Contents

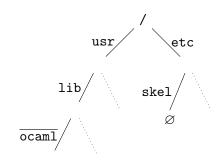
1. Description of file systems

- Unix file system
- Static description
- Directory update

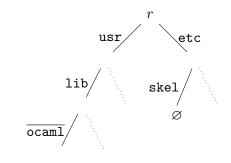
- 2. Tree constraints
 - Definitions
 - Basic constraints
 - Existential and first order constraints

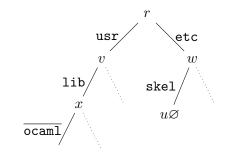


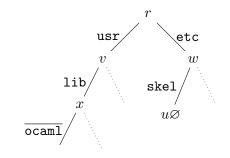
 $r \exists u, v, x, w. \quad \frac{r[\texttt{usr}]v \wedge v[\texttt{lib}]x \wedge x[\texttt{ocaml}] \uparrow}{\wedge r[\texttt{etc}]w \wedge w[\texttt{skel}]u \wedge u\varnothing}$



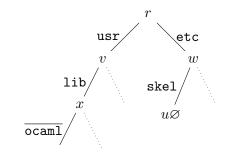
 $\begin{array}{c|c} r & \exists u, v, x, w. & r[\texttt{usr}]v \wedge v[\texttt{lib}]x \wedge x[\texttt{ocaml}] \uparrow \\ & \wedge r[\texttt{etc}]w \wedge w[\texttt{skel}]u \wedge u\varnothing \end{array}$



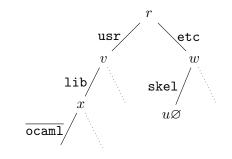




$$r \mid \exists u, v, x, w. \quad \begin{array}{c} r[\mathtt{usr}]v \wedge v[\mathtt{lib}]x \wedge x[\mathtt{ocaml}] \uparrow \\ \wedge r[\mathtt{etc}]w \wedge w[\mathtt{skel}]u \wedge u arnothing \end{array}$$



$$r \mid \exists u, v, x, w. \quad \begin{aligned} r[\texttt{usr}]v \wedge v[\texttt{lib}]x \wedge x[\texttt{ocaml}] \uparrow \\ & \wedge r[\texttt{etc}]w \wedge w[\texttt{skel}]u \wedge u \varnothing \end{aligned}$$



$$r \mid \exists u, v, x, w. \quad \begin{aligned} r[\texttt{usr}]v \wedge v[\texttt{lib}]x \wedge x[\texttt{ocaml}] \uparrow \\ & \wedge r[\texttt{etc}]w \wedge w[\texttt{skel}]u \wedge u \varnothing \end{aligned}$$

Table of Contents

1. Description of file systems

- Unix file system
- Static description
- Directory update

2. Tree constraints

- Definitions
- Basic constraints
- Existential and first order constraints



We want something like:

 $r' = r[\texttt{usr} \to v'] \land v' = v[\texttt{lib} \to x'] \land x' = x[\texttt{ocaml} \to y'] \land y' \varnothing$



We want something like:

 $r' = r[\texttt{usr} \to v'] \land v' = v[\texttt{lib} \to x'] \land x' = x[\texttt{ocaml} \to y'] \land y' \varnothing$



We want something like:

 $r' = r[\texttt{usr} \to v'] \land v' = v[\texttt{lib} \to x'] \land x' = x[\texttt{ocaml} \to y'] \land y' \varnothing$



We want something like:

$$r' = r[\texttt{usr} \to v'] \land v' = v[\texttt{lib} \to x'] \land x' = x[\texttt{ocaml} \to y'] \land y' \varnothing$$

Er.. is that really what we want?

Asymmetric:

$$y = x[f \to v]$$

• Makes it hard to eliminate variables:

$$y = x[f \to v] \land z = x[g \to w]$$

• Contains in fact two pieces of information:

ullet "y and x are different in f, identical everywhere else" :

 $y \sim_f x$

• "y points to v through f"

Er.. is that really what we want?

Asymmetric:

$$y = x[f \to v]$$

• Makes it hard to eliminate variables:

$$y = x[f \to v] \wedge z = x[g \to w]$$

Contains in fact two pieces of information:
 "y and x are different in f, identical everywhere else"

 $y \sim_f x$

• "y points to v through f"

Er.. is that really what we want?

Asymmetric:

$$y = x[f \to v]$$

• Makes it hard to eliminate variables:

$$y = x[f \to v] \land z = x[g \to w]$$

Contains in fact two pieces of information:

• "y and x are different in f, identical everywhere else"

• "y points to v through f"

Er.. is that really what we want?

Asymmetric:

$$y = x[f \to v]$$

Makes it hard to eliminate variables:

$$y = x[f \to v] \wedge z = x[g \to w]$$

- Contains in fact two pieces of information:
 - "y and x are different in f, identical everywhere else":

$$y \sim_f x$$

```
• "y points to v through f"
```

Er.. is that really what we want?

Asymmetric:

$$y = x[f \to v]$$

Makes it hard to eliminate variables:

$$y = x[f \to v] \wedge z = x[g \to w]$$

- Contains in fact two pieces of information:
 - "y and x are different in f, identical everywhere else":

$$y \sim_f x$$

• "y points to v through f":

Er.. is that really what we want?

Asymmetric:

$$y = x[f \to v]$$

Makes it hard to eliminate variables:

$$y = x[f \to v] \land z = x[g \to w]$$

- Contains in fact two pieces of information:
 - "y and x are different in f, identical everywhere else":

• "y points to v through f":

y[f]v

Er.. is that really what we want?

Asymmetric:

$$y = x[f \to v]$$

Makes it hard to eliminate variables:

$$y = x[f \to v] \wedge z = x[g \to w]$$

- Contains in fact two pieces of information:
 - "y and x are different in f, identical everywhere else":

$$y \sim_f x$$

• "y points to v through f":

y[f]v

Nicolas Jeannerod	Nico	las .	Jeannero	d
-------------------	------	-------	----------	---

Much better

Allows to express the update:

$$y = x[f \to v] := y \sim_f x \land y[f]v$$

• Symmetric:

$$y \sim_f x \Longleftrightarrow x \sim_f y$$

• Transitive:

 $y \sim_f x \wedge z \sim_f x \Longrightarrow y \sim_f z$

Much better

• Allows to express the update:

$$y = x[f \to v] := y \sim_f x \land y[f]v$$

• Symmetric:

$$y \sim_f x \Longleftrightarrow x \sim_f y$$

• Transitive:

$$y \sim_f x \wedge z \sim_f x \Longrightarrow y \sim_f z$$

Much better

• Allows to express the update:

$$y = x[f \to v] := y \sim_f x \land y[f]v$$

• Symmetric:

$$y \sim_f x \Longleftrightarrow x \sim_f y$$

• Transitive:

$$y \sim_f x \wedge z \sim_f x \Longrightarrow y \sim_f z$$

Table of Contents

1. Description of file systems

- Unix file system
- Static description
- Directory update

2. Tree constraints

Definitions

- Basic constraints
- Existential and first order constraints

Constraints

- \mathcal{K} finite set; $Dir \in \mathcal{K}$; \mathcal{F} infinite set
- \bullet Finite trees labelled with ${\cal K}$ on nodes and ${\cal F}$ on edges
- x, y variables; $K \in \mathcal{K}$, $f \in \mathcal{F}$, $F \subseteq \mathcal{F}$

Equality		K(x)	Kind
Feature	xfy	$xf\uparrow$	Absence
Fence	xF		Similarity

- Composed with \neg , \land , \lor , $\exists x$, $\forall x$
- No quantification on kinds and features
- Wanted: (un)satisfiability of these constraints
- Bonus point for incremental procedures

Constraints

- \mathcal{K} finite set; $Dir \in \mathcal{K}$; \mathcal{F} infinite set
- \bullet Finite trees labelled with ${\cal K}$ on nodes and ${\cal F}$ on edges
- x, y variables; $K \in \mathcal{K}$, $f \in \mathcal{F}$, $F \subseteq \mathcal{F}$

Equality	$x \doteq y$	K(x)	Kind
Feature	xfy	$xf\uparrow$	Absence
Fence	xF	$x \sim_F y$	Similarity

- Composed with \neg , \land , \lor , $\exists x$, $\forall x$
- No quantification on kinds and features
- Wanted: (un)satisfiability of these constraints
- Bonus point for incremental procedures

Constraints

- \mathcal{K} finite set; $Dir \in \mathcal{K}$; \mathcal{F} infinite set
- \bullet Finite trees labelled with ${\cal K}$ on nodes and ${\cal F}$ on edges
- x, y variables; $K \in \mathcal{K}$, $f \in \mathcal{F}$, $F \subseteq \mathcal{F}$

Equality	$x \doteq y$	K(x)	Kind
Feature	xfy	$xf\uparrow$	Absence
Fence	xF	$x \sim_F y$	Similarity

• Composed with \neg , \land , \lor , $\exists x$, $\forall x$

- No quantification on kinds and features
- Wanted: (un)satisfiability of these constraints
- Bonus point for incremental procedures

Constraints

- \mathcal{K} finite set; $Dir \in \mathcal{K}$; \mathcal{F} infinite set
- \bullet Finite trees labelled with ${\cal K}$ on nodes and ${\cal F}$ on edges
- x, y variables; $K \in \mathcal{K}$, $f \in \mathcal{F}$, $F \subseteq \mathcal{F}$

Equality	$x \doteq y$	K(x)	Kind
Feature	xfy	$xf\uparrow$	Absence
Fence	xF	$x \sim_F y$	Similarity

- Composed with \neg , \land , \lor , $\exists x$, $\forall x$
- No quantification on kinds and features
- Wanted: (un)satisfiability of these constraints
- Bonus point for incremental procedures

Constraints

- \mathcal{K} finite set; $Dir \in \mathcal{K}$; \mathcal{F} infinite set
- \bullet Finite trees labelled with ${\cal K}$ on nodes and ${\cal F}$ on edges
- x, y variables; $K \in \mathcal{K}$, $f \in \mathcal{F}$, $F \subseteq \mathcal{F}$

Equality	$x \doteq y$	K(x)	Kind
Feature	xfy	$xf\uparrow$	Absence
Fence	xF	$x \sim_F y$	Similarity

- Composed with \neg , \land , \lor , $\exists x$, $\forall x$
- No quantification on kinds and features
- Wanted: (un)satisfiability of these constraints
- Bonus point for incremental procedures

Game plan

1. Write a system of rewriting rules;

Game plan

- 1. Write a system of rewriting rules;
- 2. Prove that the system terminates (help it if needed);
- 3. Prove that the rules respect equivalences:

Lemma

If ϕ reduces to ψ , then $\models \phi \leftrightarrow \psi$.

4. Prove nice properties on the normal forms:

Lemma

If ϕ is in normal form, then it is either satisfiable or \perp .

Game plan

- 1. Write a system of rewriting rules;
- 2. Prove that the system terminates (help it if needed);
- 3. Prove that the rules respect equivalences:

Lemma If ϕ reduces to ψ , then $\models \phi \leftrightarrow \psi$.

4. Prove nice properties on the normal forms:

Lemma

If ϕ is in normal form, then it is either satisfiable or \bot .

Game plan

- 1. Write a system of rewriting rules;
- 2. Prove that the system terminates (help it if needed);
- 3. Prove that the rules respect equivalences:

Lemma If ϕ reduces to ψ , then $\models \phi \leftrightarrow \psi$.

4. Prove nice properties on the normal forms:

Lemma

If ϕ is in normal form, then it is either satisfiable or $\bot.$

Table of Contents

1. Description of file systems

- Unix file system
- Static description
- Directory update

2. Tree constraints

- Definitions
- Basic constraints
- Existential and first order constraints

Basic constraints: conjunction of positive atoms.

 $\frac{SIMPL-FEATS}{xfy \wedge xfz}$ $\frac{xfy \wedge y \doteq z}{xfy \wedge y \doteq z}$





INTRO-SIM-SIMS $x \sim_F y \wedge y \sim_G z$

 $x \sim_F y \wedge y \sim_G z \wedge x \sim_{(F \cup G)} z$

Basic constraints: conjunction of positive atoms.

SIMPL-FEATS C-FEAT-ABS $\frac{xfy \wedge xfz}{xfy \wedge y \doteq z}$

$$rac{xfy \wedge x}{\perp}$$

INTRO-FEAT-SIM

$$\frac{x \sim_F y \wedge xfz}{x \sim_F y \wedge xfz \wedge yfz} f \notin F$$

 $f\uparrow$

Basic constraints: conjunction of positive atoms.

 $\begin{array}{c} \text{SIMPL-FEATS} \\ \hline xfy \wedge xfz \\ \hline xfy \wedge y \doteq z \end{array} \qquad \begin{array}{c} \text{C-FEAT-ABS} \\ \hline xfy \wedge xf\uparrow \\ \bot \end{array}$

INTRO-FEAT-SIM

$$\frac{x \sim_F y \wedge xfz}{x \sim_F y \wedge xfz \wedge yfz} f \notin F$$

INTRO-SIM-SIMS $x \sim_F y \wedge y \sim_G z$

 $x \sim_F y \wedge y \sim_G z \wedge x \sim_{(F \cup G)} z$

Basic constraints: conjunction of positive atoms.

 $\begin{array}{c} \text{SIMPL-FEATS} \\ \hline xfy \wedge xfz \\ \hline xfy \wedge y \doteq z \end{array} \qquad \begin{array}{c} \text{C-FEAT-ABS} \\ \hline xfy \wedge xf\uparrow \\ \bot \end{array}$

$$\frac{x \sim_F y \wedge xfz}{x \sim_F y \wedge xfz \wedge yfz} f \notin F$$

 $\frac{x \sim_F y \land y \sim_G z}{x \sim_F y \land y \sim_G z \land x \sim_{(F \cup G)} z}$

Basic constraints: conjunction of positive atoms

- Equality: rewritten
- Kind: Static "positive" information
- Feature: Static "positive" information
- Absence: Static "negative" information
- Fence: Static "negative" information
- Similarity: **Dynamic** information

 $r[{ t usr}]v \wedge v[{ t lib}]x \wedge x[{ t ocaml}] \wedge v[{ t skel}]u \wedge u arnothing \ \wedge \ldots$



Basic constraints: conjunction of positive atoms

• Equality: rewritten

- Kind: Static "positive" information
- Feature: Static "positive" information
- Absence: Static "negative" information
- Fence: Static "negative" information
- Similarity: Dynamic information

 $r[{ t usr}]v \wedge v[{ t lib}]x \wedge x[{ t ocaml}] \cap v[{ t tcaml}] \wedge v[{ t skel}]u \wedge u \otimes v[$



Basic constraints: conjunction of positive atoms

• Equality: rewritten

- Kind: Static "positive" information
- Feature: Static "positive" information
- Absence: Static "negative" information
- Fence: Static "negative" information
- Similarity: Dynamic information

 $r[{ t usr}]v \wedge v[{ t lib}]x \wedge x[{ t ocaml}] \cap v[{ t ctarl}] \wedge v[{ t skel}]u \wedge u \otimes v[{ t skel}]u \wedge u \otimes v[{ t ocaml}]$



Basic constraints: conjunction of positive atoms

- Equality: rewritten
- Kind: Static "positive" information
- Feature: Static "positive" information
- Absence: Static "negative" information
- Fence: Static "negative" information
- Similarity: Dynamic information

 $r[{ t usr}]v \wedge v[{ t lib}]x \wedge x[{ t ocaml}] \cap v[{ t transformula}] \wedge v[{ t skel}]u \wedge u \otimes v[{ t skel}]u \wedge$



Journées PPS

Basic constraints: conjunction of positive atoms

- Equality: rewritten
- Kind: Static "positive" information
- Feature: Static "positive" information
- Absence: Static "negative" information
- Fence: Static "negative" information
- Similarity: Dynamic information

 $r[{ t usr}]v \wedge v[{ t lib}]x \wedge x[{ t ocaml}] \cap v[{ t tcl}]v \wedge w[{ t skel}]u \wedge u \oslash \dots$



Basic constraints: conjunction of positive atoms

- Equality: rewritten
- Kind: Static "positive" information
- Feature: Static "positive" information
- Absence: Static "negative" information
- Fence: Static "negative" information
- Similarity: Dynamic information

```
r[\texttt{usr}]v \wedge v[\texttt{lib}]x \wedge x[\texttt{ocaml}] \uparrow \land r[\texttt{etc}]w \wedge w[\texttt{skel}]u \wedge u\varnothing \land \ldots
```



Basic constraints: conjunction of positive atoms

- Equality: rewritten
- Kind: Static "positive" information
- Feature: Static "positive" information
- Absence: Static "negative" information
- Fence: Static "negative" information
- Similarity: Dynamic information

```
r[\texttt{usr}]v \wedge v[\texttt{lib}]x \wedge x[\texttt{ocaml}] \uparrow \land r[\texttt{etc}]w \wedge w[\texttt{skel}]u \wedge u arnothing \land \ldots
```



Basic constraints: conjunction of positive atoms

- Equality: rewritten
- Kind: Static "positive" information
- Feature: Static "positive" information
- Absence: Static "negative" information
- Fence: Static "negative" information
- Similarity: Dynamic information

```
r[\texttt{usr}]v \wedge v[\texttt{lib}]x \wedge x[\texttt{ocaml}] \uparrow \land r[\texttt{etc}]w \wedge w[\texttt{skel}]u \wedge u arnothing \land \ldots
```

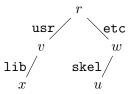


Table of Contents

1. Description of file systems

- Unix file system
- Static description
- Directory update

2. Tree constraints

- Definitions
- Basic constraints
- Existential and first order constraints

- $\neg xF$: there exist a $g \notin F$ such that $xg\downarrow$;
- $x \not\sim_F y$: there exist a $g \notin F$ such that $x \neq_g y$;



Repl-NAbs	
$\neg xf\uparrow$	
$\exists z.xfz$	

 $\frac{xF \wedge x \not\sim_G y}{xF \wedge \neg xG} \quad F \subseteq G$

- $\neg xF$: there exist a $g \notin F$ such that $xg\downarrow$;
- $x \not\sim_F y$: there exist a $g \notin F$ such that $x \neq_g y$;

$$\frac{\nabla K(x)}{\bigvee_{L \in \mathcal{K}} L(x)}$$

 $\frac{REPL-NABS}{\neg xf\uparrow}$ $\frac{\neg xf\uparrow}{\exists z.xfz}$

 $\frac{xF \wedge x \not\sim_G y}{xF \wedge \neg xG} \quad F \subseteq G$

- $\neg xF$: there exist a $g \notin F$ such that $xg\downarrow$;
- $x \not\sim_F y$: there exist a $g \notin F$ such that $x \neq_g y$;

$$\frac{\mathbb{R}_{EPL-NKINE}}{\bigvee_{L \in \mathcal{K}} L(x)}$$

Repl-NAbs		
$\neg xf\uparrow$	$xF \wedge x \not\sim_G y$	$F \subset C$
$\exists z.xfz$	$xF \wedge \neg xG$	$\Gamma \subseteq 0$

- $\neg xF$: there exist a $g \notin F$ such that $xg\downarrow$;
- $x \not\sim_F y$: there exist a $g \notin F$ such that $x \neq_g y$;

$$\frac{\mathbb{R}_{EPL-NKINE}}{\bigvee_{L \in \mathcal{K}} L(x)}$$

Repl-NAbs	Repl-NSim-NFence	
$\neg xf\uparrow$	$xF \wedge x \not\sim_G y$	$E \subset C$
$\exists z.xfz$	$\overline{xF \wedge \neg xG}$	$F \subseteq G$

Quantifier elimination

• Goal: be able to change an existentially quantified block into a universally quantified one.

$$\left(\exists \vec{X}. \bigwedge \dots \right) \leftrightarrow \left(\forall \vec{Y}. \bigvee \bigwedge \dots \right)$$

• Special rules:

 $\frac{\exists x. \exists \vec{X}. (yfx \land \phi(x, \vec{X}))}{yf \downarrow \land \forall x. \exists \vec{X}. (yfx \to \phi(x, \vec{X}))}$

Quantifier elimination

• Goal: be able to change an existentially quantified block into a universally quantified one.

$$\left(\exists \vec{X}. \bigwedge \dots \right) \leftrightarrow \left(\forall \vec{Y}. \bigvee \bigwedge \dots \right)$$

• Special rules:

$$\frac{\exists x. \exists \vec{X}. (yfx \land \phi(x, \vec{X}))}{yf \downarrow \land \forall x. \exists \vec{X}. (yfx \to \phi(x, \vec{X}))}$$

Lemma of 31 August

Lemma (31 August)

Let ϕ be a conjunction of the form:

$$\phi(\vec{X},\vec{Y}) = \left(\bigwedge \textit{stuff about } \vec{X}\right) \wedge \psi(\vec{Y})$$

in normal form for our system. Then we have:

$$\models \forall \vec{Y}. \left(\exists \vec{X}. \ \phi(\vec{X}, \vec{Y}) \right) \leftrightarrow \psi(\vec{Y})$$

• The system propagates all the useful information

• We can just remove what we don't need!

Lemma of 31 August

Lemma (31 August)

Let ϕ be a conjunction of the form:

$$\phi(\vec{X},\vec{Y}) = \left(\bigwedge \textit{stuff about } \vec{X}\right) \wedge \psi(\vec{Y})$$

in normal form for our system. Then we have:

$$\models \forall \vec{Y}. \left(\exists \vec{X}. \ \phi(\vec{X}, \vec{Y}) \right) \leftrightarrow \psi(\vec{Y})$$

- The system propagates all the useful information
- We can just remove what we don't need!

Thank you for your attention!

Recap':

- Need constraints on graphs to represent relations on file systems;
- Extend "feature trees" with x ∼_F y ("x and y are the same, except maybe for the features in F");
- Use a system of rewrite rules whose normal forms have nice properties.

Future work:

- Cleanup, formalise in a technical report;
- Add inodes, permissions, timestamps, etc.
- Implement an efficient version for the existential subset.