The primitivity of operators in the algebra of binary relations under conjunctions of containments

Dimitri Surinx ¹  Jan Van den Bussche ¹  Dirk Van Gucht ²

¹Hasselt University ²Indiana University
Algebra of binary relations

Operations on binary relations over some domain $V$ [Peirce, Schröder, Tarski]

\[
\begin{align*}
\text{id} &= \{(x, x) \mid x \in V\} \\
 r^{-1} &= \{(x, y) \mid (y, x) \in r\} \\
 r \circ s &= \{(x, y) \mid \exists z : (x, z) \in r \land (z, y) \in s\} \\
 r \cup s &= \{(x, y) \mid (x, y) \in r \lor (x, y) \in s\} \\
 r - s &= \{(x, y) \mid (x, y) \in r \land (x, y) \not\in s\} \\
 r \cap s &= \{(x, y) \mid (x, y) \in r \land (x, y) \in s\}
\end{align*}
\]
Additional Operations

\[ \pi_1(r) = \{(x, x) \mid \exists y : (x, y) \in r\} \]  
\[ \pi_2(r) = \{(x, x) \mid \exists y : (y, x) \in r\} \]  
\[ \overline{\pi}_1(r) = \{(x, x) \mid \neg(\exists y) : (x, y) \in r\} \]  
\[ \overline{\pi}_2(r) = \{(x, x) \mid \neg(\exists y) : (y, x) \in r\} \]  

(provide tests)

(provide negative tests)

\[ all = V^2 \]

\[ di = all - id \]

\[ r^+ = \text{the transitive closure of } r \]
Fragments & Path Queries

The most basic fragment $\mathcal{N}$ is the semiring $\{\emptyset, id, \cup, \circ\}$

$\mathcal{N}(F)$ denotes $\mathcal{N}$ extended with the operations in $F$

Fix a binary relational vocabulary $\Gamma$ (Label set)

Expressions over a fragment $F$ are built from relation names in $\Gamma$ using the operations in $F$

$\Rightarrow$ Map instances of $\Gamma$ (graphs) to binary relations (*path query*)

Write $F_1 \leq F_2$ if every path query expressible in $\mathcal{N}(F_1)$ is also expressible in $\mathcal{N}(F_2)$

Well established logic: $\mathcal{N}(-1, -, di) \equiv \text{FO}^3$ (Tarski & Givant)

The primitivity of operators in the algebra of binary relations under conjunctions of containments
Examples

Consider Hobbies(person,hname), Person(pname,age)

“Output the persons that share a hobby”
- \((\text{Hobbies} \circ (\text{Hobbies}^{-1})) \circ \text{id}\)

“Output the persons that do not have any hobbies”
- \(\pi_1(\text{Hobbies}) \circ \pi_1(\text{Person})\)
Some operators can be constructed with other operators:

\[
\begin{align*}
\text{all} & \equiv \text{di} \cup \text{id} \\
\text{di} & \equiv \text{all} - \text{id} \\
e_1 \cap e_2 & \equiv e_1 - (e_1 - e_2) \\
\pi_1(e) & \equiv (e \circ e^{-1}) \cap \text{id} \equiv (e \circ \text{all}) \cap \text{id} \equiv \pi_1(\pi_1(e)) \\
\pi_2(e) & \equiv (e^{-1} \circ e) \cap \text{id} \equiv (\text{all} \circ e) \cap \text{id} \equiv \pi_2(\pi_2(e)) \\
\overline{\pi}_1(e) & \equiv \text{id} - \pi_1(e) \\
\overline{\pi}_2(e) & \equiv \text{id} - \pi_2(e)
\end{align*}
\]

⚠️ Operator can be expressed in \( \mathcal{N}(F) \) without belonging to \( F \)

The primitivity of operators in the algebra of binary relations under conjunctions of containments
Define $\overline{F}$ as the smallest superset of $F$ that satisfy the *completion axioms*:

- $di \in \overline{F}$, then $all \in \overline{F}$
- $\pi \in \overline{F}$, then $\pi \in \overline{F}$
- $\cap \in \overline{F}$ and $di \in \overline{F}$, then $\pi \in \overline{F}$
- $\cap \in \overline{F}$ and $^{-1} \in \overline{F}$, then $\pi \in \overline{F}$
- $- \in \overline{F}$ and $\pi \in \overline{F}$, then $\bar{\pi} \in \overline{F}$
- $- \in \overline{F}$, then $\cap \in \overline{F}$

Example: $\overline{\{all, -\}} = \{di, all, \cap, -, \pi, \bar{\pi}\}$

Clearly, $\mathcal{N}(F) \equiv \mathcal{N}(\overline{F})$
An operator \( g \) is \emph{primitive} (for path queries) if for every fragment \( F \) with \( g \not\in \overline{F} \): \( \{g\} \not\subseteq F \)

\begin{quote}
\textbf{Theorem ([Fletcher et al., 2011])}\\
\textit{Every operator is primitive for path queries}
\end{quote}

\( \Rightarrow \) The completion axioms are complete for the purpose of comparing fragments
Nonemptiness

A Boolean query maps graphs to yes/no

Typically expressed by nonemptiness expressions $e \neq \emptyset$
Example: $R^2 - R \neq \emptyset$ (nontransitivity)

Define $F \neq \emptyset$ as the set of boolean queries expressible in the form $e \neq \emptyset$
with $e \in \mathcal{N}(F)$

The primitivity of operators in the algebra of binary relations under conjunctions of containments
Primitivity for nonemptiness

Converse and Transitive closure are not primitive for nonemptiness:

Theorem ([Fletcher et al., 2011, 2013])

If $F$ has projection, but neither intersection nor transitive closure, then

$$F \neq \emptyset \subseteq F - \{-1\} \neq \emptyset$$

If $F \subseteq \{\pi, di, +\}$ and $\Gamma$ only contains one edge label, then

$$F \neq \emptyset \subseteq F - \{+\} \neq \emptyset$$

Examples:

- $R^2 \circ R^{-1} \circ R^3 \neq \emptyset \equiv \pi_1(R^2 \circ \pi_2(\pi_1(R^3) \circ R)) \neq \emptyset$
- $\pi_1(R^3) \circ R^+ \circ di \circ \pi_2(R) \circ R^2 \neq \emptyset \equiv \pi_1(R^3) \circ (R \cup R^2) \circ di \circ \pi_2(R) \circ R^2 \neq \emptyset$
Conjunctions of Containments

Nonemptiness is not the only way for expressing boolean queries
- Example: consider $e = \emptyset$ statements

Natural to express properties as finite conjunctions of containment statements
- A containment statement $e_1 \subseteq e_2$ holds on $G$ if $e_1(G) \subseteq e_2(G)$

Examples:
- Total order: $R^2 \subseteq R \land id \subseteq R \land R \cap R^{-1} \subseteq id \land all \subseteq R \cup R^{-1}$
- Function: $R^{-1} \circ R \subseteq id$
- Connectivity: $all \subseteq (R \cup R^{-1})^+$

The primitivity of operators in the algebra of binary relations under conjunctions of containments
Main result

For a fragment $F$ define $F^\subseteq$ as the boolean queries expressible by finite conjunctions of containment statements in $\mathcal{N}(F)$

**Theorem**

*Every operator is primitive for finite conjunctions of containments*

In other words, if $g \not\in \overline{F}$ then $\{g\}^\subseteq \not\subseteq F^\subseteq$

$\Rightarrow$ The completion axioms are complete for comparing fragments under finite conjunctions of containments
Fragments with Difference

Notice that $e_1 \subseteq e_2 \land e_3 \subseteq e_4 \equiv e_1 - e_2 \cup e_3 - e_4 = \emptyset$

$\Rightarrow$ emptiness $\equiv$ containment when difference is present

When difference is present inexpressibility sometimes reduces to inexpressibility for (non)emptiness

- Example: $\{\pi\} \subseteq \not\subseteq \{-, +\} \subseteq$ reduces to $\{\pi\} = \emptyset \not\subseteq \{-, +\} = \emptyset$
Primitivity of Projection

The only fragment of interest is \( \{di, -1, +\} \)

“Paths of length 2 only start in loops” is not expressible in \( \{di, -1, +\} \subseteq \)

- Expressed by \( R \circ \pi_1(R) \subseteq id \) in \( \{\pi\} \subseteq \)

We prove that queries in \( \{di, -1, +\} \subseteq \) cannot be false on \( K_3 \), true on \( H \) and true on \( id_3 \) simultaneously
The only fragment of interest is \(\{d_i, \cap, -1, +\}\).

“Every node is at most one step away from a sink node” is not expressible in \(\{d_i, \cap, -1, +\}\)⊆

- Expressed by \(\pi_1(R) \subseteq \pi_1(R \circ \pi_1(R))\) in \(\{\pi\}\)⊆

We prove that queries in \(\{d_i, \cap, -1, +\}\)⊆ cannot be false on \(K_3\) and true on \(H\) simultaneously.
Primitivity of Difference

The only fragment of interest is \( \{ di, \overline{\pi}, -1, \cap, + \} \)

The query \( id \subseteq R^2 \circ (R^2 - R) \circ R^2 \) is not expressible in \( \{ di, \overline{\pi}, -1, \cap, + \} \subseteq \)

We prove that queries in \( \{ di, \overline{\pi}, -1, \cap, + \} \subseteq \) cannot be false on \( K_3 \) and true on \( B \) simultaneously.

\[ K_3 \]

\[ B \]
The only fragment of interest is $\{di, \overline{\pi}, -1, +\}$

"The only transitive edges are loops" is not expressible in $\{di, \overline{\pi}, -1, +\} \subseteq \{\cap\}$

- Expressed by $R^2 \cap R \subseteq id$ in $\{\cap\} \subseteq \{\cap\}$

We prove that queries $\{di, \overline{\pi}, -1, +\} \subseteq \{\cap\}$ cannot be false on $K_3$, true on $id_3$ and true on $\ell_2$ simultaneously.
Primitivity of Transitive Closure

The fragment of interest is $\{-1, di, -\}$

“There is a cycle through every node” is not expressible in $\{-1, di, -\} \subseteq \{id \subseteq R^+ \in \{+\} \subseteq \}$

⚠️ In contrast with the collapses of $+$ under nonemptiness!

We show that the families of graphs $G_1^\ell$ and $G_2^\ell$ are not distinguishable in FO

The primitivity of operators in the algebra of binary relations under conjunctions of containments
The fragment of interest is \( \{-1, -, +\} \)

“The graph is complete” is not expressible in \( \{-1, -, +\} \subseteq \)

- Expressed by \( all \subseteq R \) in \( \{all\} \subseteq \)

We prove that queries in \( \{-1, -, +\} \subseteq \) cannot be true on a single self-loop and false on two disjoint self-loops simultaneously

\[
\begin{align*}
\bullet & \\
\text{id}_1 & \quad \text{id}_2
\end{align*}
\]
The only fragment of interest is \( \{-1, \cap, all, \bar{\pi}, +\} \)

“There is only one node” is not expressible in \( \{-1, \cap, all, \bar{\pi}, +\} \subseteq \)

- Expressed by \( di \subseteq \emptyset \) in \( \{di\} \subseteq \)

We prove that queries in \( \{-1, \cap, all, \bar{\pi}, +\} \subseteq \) cannot be true on \( id_1 \) and false on \( K_3 \) simultaneously

The primitivity of operators in the algebra of binary relations under conjunctions of containments
Primitivity of Converse

The fragment of interest is \( \{d_i, -, +\} \)

“If there is a path \( a \rightarrow b \rightarrow c \leftarrow d \rightarrow e \) with \( d \neq b \) then \( a \rightarrow e \)” is not expressible in \( \{d_i, -, +\} \)

- Expressed by \( R^2 \circ R^{-1} \circ R \subseteq R^2 \cup R \) in \( \{-1\} \subseteq \)

⚠️ In contrast with the collapses of \(-1\) under nonemptiness!

We prove that queries in \( \{d_i, -, +\} \subseteq \) cannot be true on \( G_1 \) and false on \( G_2 \) simultaneously
Directions for further research

- All of our separating queries do not use conjunction. When does conjunction actually add expressive power?
- Compare the expressive power of nonemptiness and containment:
  
  Largely completed this work, one open question remains:

  \[ \{ \pi \} \neq \emptyset \subseteq \{ di, -1, + \} \subseteq \]

- Another interesting derived operator: Residuations [Pratt. Origins of calculus of binary relations]
  
  \( C / B \) is the maximal \( X \) such that

  \[ X \circ B \subseteq C \]

  \( B \setminus C \) is the maximal \( X \) such that

  \[ B \circ X \subseteq C \]