

Semantic Web – Tutorial #10

Isabelle Kuhlmann

Institute for Web Science & Technologies
University of Koblenz-Landau

July 8, 2021 | Assignment 9

1: Ontology-based Data Access

Consider the following Global Schema \mathcal{G} :

$$\mathcal{G} = \left\{ \begin{array}{ll} \text{Book} & \sqsubseteq \exists \text{hasTitle}.\top \sqcap \exists \text{publishedInYear}.\top \sqcap \exists \text{hasAuthor}.\top \sqcap \exists \text{hasPrice}.\top \\ \text{Reader} & \sqsubseteq \exists \text{reads}.\text{Book} \sqcap \exists \text{hasAge}.\top \\ \text{Novel} & \sqsubseteq \text{Book} \end{array} \right\}$$

Also consider the two data sources $\mathcal{S} = \mathcal{S}_1 \mathcal{S}_2$ with:

- ▶ $\mathcal{S}_1 = \{R1(\text{bookTitle}, \text{author}, \text{year}, \text{price})\}$, which contains novels published after 1990
- ▶ $\mathcal{S}_2 = \{R2(\text{bookTitle}, \text{reader})\}$, which contains books read by people who are under the age of 30

Schema and Data Sources

$$\mathcal{G} = \left\{ \begin{array}{ll} \text{Book} & \sqsubseteq \exists \text{hasTitle}.\top \sqcap \exists \text{publishedInYear}.\top \sqcap \exists \text{hasAuthor}.\top \sqcap \exists \text{hasPrice}.\top \\ \text{Reader} & \sqsubseteq \exists \text{reads}.\text{Book} \sqcap \exists \text{hasAge}.\top \\ \text{Novel} & \sqsubseteq \text{Book} \end{array} \right\}$$

- ▶ $S_1 = \{R_1(\text{bookTitle}, \text{author}, \text{year}, \text{price})\}$, which contains novels published after 1990
- ▶ $S_2 = \{R_2(\text{bookTitle}, \text{reader})\}$, which contains books read by people who are under the age of 30

Task: Formalize GAV mappings for the global schema (\mathcal{G}) and data sources (S).

SELECT R1.bookTitle AS x FROM R1 \rightarrow x: Novel

SELECT R1.bookTitle AS x FROM R1

UNION SELECT R2.bookTitle AS x FROM R2 \rightarrow x: Book

SELECT R1.bookTitle AS x FROM R1

UNION SELECT R2.bookTitle AS x FROM R2 \rightarrow (x,x): hasTitle

SELECT R1.bookTitle AS x FROM R1,

R1.year AS y FROM R1 \rightarrow (x,y): publishedInYear

Schema and Data Sources

$\mathcal{G} =$ { Book \sqsubseteq $\exists \text{hasTitle}.\top \sqcap \exists \text{publishedInYear}.\top \sqcap \exists \text{hasAuthor}.\top \sqcap \exists \text{hasPrice}.\top$
 Reader \sqsubseteq $\exists \text{reads}.\text{Book} \sqcap \exists \text{hasAge}.\top$
 Novel \sqsubseteq Book }

▶ $S_1 = \{R1(\text{bookTitle}, \text{author}, \text{year}, \text{price})\}$, which contains novels published after 1990

▶ $S_2 = \{R2(\text{bookTitle}, \text{reader})\}$, which contains books read by people who are under the age of 30

SELECT R1.bookTitle AS x FROM R1,
R1.author AS y FROM R1 \rightarrow (x,y): hasAuthor

SELECT R1.bookTitle AS x FROM R1,
R1.price AS y FROM R1 \rightarrow (x,y): hasPrice

SELECT R2.reader AS z FROM R2 \rightarrow z: Reader

SELECT R2.bookTitle AS x FROM R2,
R2.reader AS z FROM R2 \rightarrow (z, x): reads

SELECT R2.bookTitle AS x FROM R2,
R2.reader AS z FROM R2 \rightarrow (z, g(z)): hasAge

Schema and Data Sources

$\mathcal{G} =$ { Book \sqsubseteq $\exists \text{hasTitle}.\top \sqcap \exists \text{publishedInYear}.\top \sqcap \exists \text{hasAuthor}.\top \sqcap \exists \text{hasPrice}.\top$
 Reader \sqsubseteq $\exists \text{reads}.\text{Book} \sqcap \exists \text{hasAge}.\top$
 Novel \sqsubseteq Book }

▶ $\mathcal{S}_1 = \{R1(\text{bookTitle}, \text{author}, \text{year}, \text{price})\}$, which contains novels published after 1990

▶ $\mathcal{S}_2 = \{R2(\text{bookTitle}, \text{reader})\}$, which contains books read by people who are under the age of 30

Task: Formalize LAV mappings for the global schema (\mathcal{G}) and data sources (\mathcal{S}).

SELECT R1.bookTitle AS x, R1.author AS y, R1.year AS z, R1.price AS w FROM R1

→ x: Novel \wedge
 (x,x): hasTitle \wedge
 (x,y): hasAuthor \wedge
 (x,z): publishedInYear \wedge
 (x,w): hasPrice \wedge
 z > 1990

Schema and Data Sources

$\mathcal{G} =$ { Book \sqsubseteq $\exists \text{hasTitle}.\top \sqcap \exists \text{publishedInYear}.\top \sqcap \exists \text{hasAuthor}.\top \sqcap \exists \text{hasPrice}.\top$
 Reader \sqsubseteq $\exists \text{reads}.\text{Book} \sqcap \exists \text{hasAge}.\top$
 Novel \sqsubseteq Book }

- ▶ $S_1 = \{R1(\text{bookTitle}, \text{author}, \text{year}, \text{price})\}$, which contains novels published after 1990
- ▶ $S_2 = \{R2(\text{bookTitle}, \text{reader})\}$, which contains books read by people who are under the age of 30

SELECT R2.bookTitle AS x, R2.reader AS y FROM R2

→ $\exists z:$

x: Book \wedge

y: Reader \wedge

(y,x): reads \wedge

(y,z): hasAge \wedge

z < 30

Schema and Data Sources

$$\mathcal{G} = \left\{ \begin{array}{ll} \text{Book} & \sqsubseteq \exists \text{hasTitle}.\top \sqcap \exists \text{publishedInYear}.\top \sqcap \exists \text{hasAuthor}.\top \sqcap \exists \text{hasPrice}.\top \\ \text{Reader} & \sqsubseteq \exists \text{reads}.\text{Book} \sqcap \exists \text{hasAge}.\top \\ \text{Novel} & \sqsubseteq \text{Book} \end{array} \right\}$$

- ▶ $\mathcal{S}_1 = \{R1(\text{bookTitle}, \text{author}, \text{year}, \text{price})\}$, which contains novels published after 1990
- ▶ $\mathcal{S}_2 = \{R2(\text{bookTitle}, \text{reader})\}$, which contains books read by people who are under the age of 30

Task: Formalize the following query:

Show readers who read books by Goethe.

$q1 = x : \text{Reader} \wedge y : \text{Book} \wedge (x,y) : \text{reads} \wedge (y, \text{Goethe}) : \text{hasAuthor}$

Schema and Data Sources

$$\mathcal{G} = \left\{ \begin{array}{ll} \text{Book} & \sqsubseteq \exists \text{hasTitle}.\top \sqcap \exists \text{publishedInYear}.\top \sqcap \exists \text{hasAuthor}.\top \sqcap \exists \text{hasPrice}.\top \\ \text{Reader} & \sqsubseteq \exists \text{reads.Book} \sqcap \exists \text{hasAge}.\top \\ \text{Novel} & \sqsubseteq \text{Book} \end{array} \right\}$$

- ▶ $\mathcal{S}_1 = \{R1(\text{bookTitle}, \text{author}, \text{year}, \text{price})\}$, which contains novels published after 1990
- ▶ $\mathcal{S}_2 = \{R2(\text{bookTitle}, \text{reader})\}$, which contains books read by people who are under the age of 30

Task: Formalize the following query:

Show all novels that cost more than \$30 and that are read by people under the age of 20.

$$q_2 = x : \text{Novel} \wedge (x,y) : \text{hasPrice} \wedge y > 30 \wedge z : \text{Reader} \wedge (z,x) : \text{reads} \wedge (z,w) : \text{hasAge} \wedge w < 20$$

Schema and Data Sources

$$\mathcal{G} = \left\{ \begin{array}{ll} \text{Book} & \sqsubseteq \exists \text{hasTitle}.\top \sqcap \exists \text{publishedInYear}.\top \sqcap \exists \text{hasAuthor}.\top \sqcap \exists \text{hasPrice}.\top \\ \text{Reader} & \sqsubseteq \exists \text{reads}.\text{Book} \sqcap \exists \text{hasAge}.\top \\ \text{Novel} & \sqsubseteq \text{Book} \end{array} \right\}$$

- ▶ $\mathcal{S}_1 = \{R1(\text{bookTitle}, \text{author}, \text{year}, \text{price})\}$, which contains novels published after 1990
- ▶ $\mathcal{S}_2 = \{R2(\text{bookTitle}, \text{reader})\}$, which contains books read by people who are under the age of 30

Task: Formalize the following query:

Show novels that were published in 1999 and that cost at least \$40.

$q3 = x : \text{Novel} \wedge (x, 1999) : \text{publishedInYear} \wedge (x, y) : \text{hasPrice} \wedge y \geq 40$

Schema and Data Sources

$\mathcal{G} =$ { Book \sqsubseteq $\exists \text{hasTitle.}\top \sqcap \exists \text{publishedInYear.}\top \sqcap \exists \text{hasAuthor.}\top \sqcap \exists \text{hasPrice.}\top$
 Reader \sqsubseteq $\exists \text{reads.Book} \sqcap \exists \text{hasAge.}\top$
 Novel \sqsubseteq Book }

▶ $S_1 = \{R1(\text{bookTitle}, \text{author}, \text{year}, \text{price})\}$, which contains novels published after 1990

▶ $S_2 = \{R2(\text{bookTitle}, \text{reader})\}$, which contains books read by people who are under the age of 30

Task: Formalize the following query:

Show readers under 30 who read books which were published before 1990.

$q_4 = x : \text{Reader} \wedge (x,y) : \text{hasAge} \wedge y < 30 \wedge z : \text{Book} \wedge (x,z) : \text{reads} \wedge (z,w) : \text{publishedInYear}$
 $\wedge w < 1990$