Assignment 4

201300035 方盛俊

Question 1. OWA and CWA

(1)

- (a) Album(Fantasy)
- (f)
 - ¬StudioAlbum(2004_Incomparable_Concert)∨¬LiveAlbum(2004_Incomparable_Concert)
- (h) \exists x.hasFriend(Jay_Chou,x)
- (i) ∃x.(hasFriend(Jay Chou,x)∧∃y.(dancesWith(x,y)∧Song(y)))
- (k) hasFriend(Vincent_Fang,Jay_Chou)

(2)

- (a) Album(Fantasy): No
- (b) StudioAlbum(The Eight Dimensions): Yes
- (c) LiveAlbum(Common Jasmin Orange): No
- (d) ¬LiveAlbum(Common Jasmin Orange): Yes
- (e) ¬EP(Secret): Yes
- (f) ¬StudioAlbum □ ¬LiveAlbum(2004_Incomparable_Concert): Yes
- (g) ¬StudioAlbum ⊔ ¬LiveAlbum(Eason_Chan): Yes
- (h) ∃hasFriend.T(Jay_Chou): Yes
- (i) \(\frac{1}{2}\) has Friend. \(\frac{1}{2}\) dances \(\text{With.Song(Jay Chou): Yes } \)
- (j) ∃hasFriend.Composer(Jay_Chou): No
- (k) \(\frac{1}{2}\) has Friend. \(\{ \text{Jay_Chou}\}\) (Vincent_Fang): No
- (I) DebutAlbum(2004_Incomparable_Concert): No
- (m) Song(Rewind): No
- (n) Singer(Jay_Chou): Yes
- (o) Singer(Jolin_Tsai): No
- (p) Lyricist(Jay_Chou): No
- (q) Composer(Jay Chou): Yes
- (r) Composer(Ta-yu_Lo): No
- (s) Police(Jay_Chou): No
- (t) Police(Jolin_Tsai): No

- (u) ¬Singer-SongWriter ⊔ ¬Police(Vincent_Fang): Yes
- (v) ¬Singer-SongWriter ⊔ ¬Police(Ta-yu Lo): yes
- (w) Singer-SongWriter(Jay_Chou): No
- (x) Singer-SongWriter(Jolin Tsai): No
- (y) ¬SongWriter(Vincent_Fang): Yes
- (z) ¬Dancer(Will Liu): Yes

(3)

- (a) Album(Fantasy): Don't know
- (b) StudioAlbum(The Eight Dimensions): Yes
- (c) LiveAlbum(Common_Jasmin_Orange): Don't know
- (d) ¬LiveAlbum(Common_Jasmin_Orange): Don't know
- (e) ¬EP(Secret): Don't know
- (f) ¬StudioAlbum ⊔ ¬LiveAlbum(2004 Incomparable Concert): Don't know
- (g) ¬StudioAlbum □ ¬LiveAlbum(Eason Chan): Don't know
- (h) ∃hasFriend.T(Jay_Chou): Yes
- (i) \(\frac{1}{2}\) has Friend. \(\frac{1}{2}\) dances \(\text{With. Song(Jay Chou): Yes } \)
- (j) ∃hasFriend.Composer(Jay Chou): Don't know
- (k) HasFriend. (Jay Chou) (Vincent Fang): Don't know
- (I) DebutAlbum(2004 Incomparable Concert): Don't know
- (m) Song(Rewind): Don't know
- (n) Singer(Jay Chou): Yes
- (o) Singer(Jolin Tsai): Don't know
- (p) Lyricist(Jay Chou): Don't know
- (q) Composer(Jay_Chou): Yes
- (r) Composer(Ta-yu Lo): Don't know
- (s) Police(Jay Chou): Don't know
- (t) Police(Jolin_Tsai): Don't know
- (u) ¬Singer-SongWriter □ ¬Police(Vincent Fang): Don't know
- (v) ¬Singer-SongWriter □ ¬Police(Ta-yu Lo): Don't know
- (w) Singer-SongWriter(Jay Chou): Don't know
- (x) Singer-SongWriter(Jolin_Tsai): Don't know
- (y) ¬SongWriter(Vincent_Fang): Don't know
- (z) ¬Dancer(Will Liu): Don't know

(4)

- (a) answer(F1(x), D_{music}) = { Jay_Chou, Eason_Chan }
- (b) answer(F2(x), D_{music}) = { Will_Liu, Black_Cat, The_Eight_Dimensions, Secret,
 Together, Ta-yu_Lo, Jolin_Tsai, Vincent_Fang, Hidden_Track, Herbalist_Manual, Jay,

Common_Jasmin_Orange, Initial_D, Elimination, Fantasy, Pearl_of_the_Orient, 2004_Incomparable_Concert, Rewind }

- (c) answer(F3(x, y), D_{music}) = { (Jay_Chou, Vincent_Fang), (Jay_Chou, Will_Liu) }
- (d) answer(F4(x), D_{music}) = { Vincent_Fang }

(5)

- (a) cetanswer(F1(x), D_{music}) = { Jay_Chou, Eason_Chan }
- (b) cetanswer(F2(x), D_{music}) = \emptyset
- (c) cetanswer(F3(x, y), D_{music}) = { (Jay_Chou, Vincent_Fang), (Jay_Chou, Will_Liu) }
- (d) cetanswer(F4(x), D_{music}) = Ø

Question 2. Querying with TBox

(1)

The certain answers in the context of D_{music} is same with Question 2. (3).

The certain answers in the context of $(\mathcal{T}, D_{\text{music}})$:

- (a) Album(Fantasy): Yes
- (b) StudioAlbum(The Eight Dimensions): Yes
- (c) LiveAlbum(Common Jasmin Orange): No
- (d) ¬LiveAlbum(Common Jasmin Orange): Yes
- (e) ¬EP(Secret): Don't know
- (f) ¬StudioAlbum ⊔ ¬LiveAlbum(2004_Incomparable_Concert): Yes
- (g) ¬StudioAlbum □ ¬LiveAlbum(Eason Chan): Yes
- (h) ∃hasFriend.⊤(Jay_Chou): Yes
- (i) \(\frac{1}{2}\) has Friend. \(\frac{1}{2}\) dances \(\text{With.Song(Jay Chou): Yes } \)
- (j) \(\frac{1}{2}\) has Friend. Composer (Jay_Chou): Don't know
- (k) ∃hasFriend.{Jay Chou}(Vincent Fang): Don't know
- (I) DebutAlbum(2004_Incomparable_Concert): No
- (m) Song(Rewind): Yes
- (n) Singer(Jay_Chou): Yes
- (o) Singer(Jolin_Tsai): Don't know
- (p) Lyricist(Jay_Chou): Yes
- (q) Composer(Jay_Chou): Yes
- (r) Composer(Ta-yu_Lo): Don't know
- (s) Police(Jay_Chou): No
- (t) Police(Jolin_Tsai): Don't know
- (u) ¬Singer-SongWriter □ ¬Police(Vincent Fang): Don't know

- (v) ¬Singer-SongWriter ⊔ ¬Police(Ta-yu_Lo): Don't know
- (w) Singer-SongWriter(Jay_Chou): Don't know
- (x) Singer-SongWriter(Jolin_Tsai): Don't know
- (y) ¬SongWriter(Vincent_Fang): Don't know
- (z) ¬Dancer(Will_Liu): Don't know

(2)

All assertions α :

- Song(Herbalist Manual)
- Song(Elimination)
- Singer(Jay Chou)
- Lyricist(Vincent_Fang)

Question 3. Computing $\mathcal{I}_{\mathcal{T},\mathcal{A}}$ in \mathcal{EL}

(1)

The initial assignment (with obvious abbreviations) is given by

```
• S(d<sub>Guitarist</sub>) = { Guitarist }
```

- S(d_{Bassist}) = { Bassist }
- S(d_{Drummer}) = { Drummer }
- S(d_{RockBand}) = { RockBand }
- S(d_{Manager}) = { Manager }
- S(d_{Employee}) = { Employee }
- S(John Lennon) = { Guitarist }
- S(Paul_McCartney) = { Bassist }
- S(Ringo_Starr) = { Drummer }
- S(Beatles) = { RockBand }
- S(Brian Epstein) = ∅
- R(managed by) = { (Beatles, Brian Epstein) }
- R(plays_for) = Ø

Update S using (simpleR):

• S(d_{Manager}) = { Manager, Employee }

Update R using (rightR):

- R(plays_for) = { (d_{Guitarist}, d_{RockBand}), (John_Lennon, d_{RockBand}), (d_{Bassist}, d_{RockBand}), (Paul_McCartney, d_{RockBand}), (d_{Drummer}, d_{RockBand}), (Ringo_Starr, d_{RockBand}) }
- R(managed_by) = { (Beatles, Brian_Epstein), (d_{RockBand}, d_{Manager}), (Beatles, d_{Manager}), (d_{Manager}, d_{Manager}) }

So we have:

- $\Delta_{\mathcal{T},\mathcal{A}}^{\mathcal{I}} = \{ d_{\text{Guitarist}}, d_{\text{Bassist}}, d_{\text{Drummer}}, d_{\text{RockBand}}, d_{\text{Manager}}, d_{\text{Employee}}, \text{John_Lennon}, \\ \text{Paul_McCartney, Ringo_Starr, Beatles, Brian_Epstein} \}$
- Guitarist $^{IT,A} = \{ d_{Guitarist}, Guitarist \}$
- Bassist^{IT}, A = { d_{Bassist}, Paul_McCartney }
- Drummer $^{IT,A} = \{ d_{Drummer}, Ringo_Starr \}$
- RockBand $^{IT,A} = \{ d_{RockBand}, Beatles \}$
- Manager $\mathcal{I}_{\mathcal{T},\mathcal{A}} = \{ d_{Manager} \}$
- Employee $^{IT,A} = \{ d_{Employee}, d_{Manager} \}$
- plays_for $^{\mathcal{IT},\mathcal{A}}$ = { (d_{Guitarist}, d_{RockBand}), (John_Lennon, d_{RockBand}), (d_{Bassist}, d_{RockBand}), (Paul_McCartney, d_{RockBand}), (d_{Drummer}, d_{RockBand}), (Ringo_Starr, d_{RockBand}) }
- managed_by $^{\mathcal{IT},\mathcal{A}}$ = { (Beatles, Brian_Epstein), (d_{RockBand}, d_{Manager}), (Beatles, d_{Manager}), (d_{Manager}, d_{Manager}) }

(2)

• ∃plays for.RockBand(John Lennon):

Because (John_Lennon, $d_{RockBand}$) \in plays_for $\mathcal{I}_{\mathcal{T},\mathcal{A}}$, $\mathcal{I}_{\mathcal{T},\mathcal{A}}$ gives the answer "Yes".

And $(\mathcal{T},\mathcal{A})$ gives the certain answer "Yes".

• ∃managed_by.Manager(Paul_McCartney)

There is no $x \in Manager^{\mathcal{IT},\mathcal{A}}$ and $(Paul_McCartney, x) \in managed_by^{\mathcal{IT},\mathcal{A}}$ so $\mathcal{I}_{\mathcal{T},\mathcal{A}}$ doesn't give the answer "Yes".

And $(\mathcal{T},\mathcal{A})$ doesn't give the certain answer "Yes".

• <code>∃plays_for.∃managed_by.Manager(Ringo_Starr)</code>

Because (Ringo_Starr, $d_{RockBand}$) \in plays_for^ $\mathcal{I}_{\mathcal{T},\mathcal{A}}$ and ($d_{RockBand}$, $d_{Manager}$) \in managed_by $\mathcal{I}_{\mathcal{T},\mathcal{A}}$, $\mathcal{I}_{\mathcal{T},\mathcal{A}}$ gives the answer "Yes".

And $(\mathcal{T},\mathcal{A})$ gives the certain answer "Yes".

```
• F(x, y) = \exists z.(plays\_for(x, z) \land plays\_for(y, z))
```

```
answer(F(x, y), \mathcal{I}_{\mathcal{T}, \mathcal{A}} = \{ (d_{Guitarist}, d_{Guitarist}), (d_{Guitarist}, John_Lennon), (d_{Guitarist}, d_{Bassist}), (d_{Guitarist}, Paul_McCartney), (d_{Guitarist}, d_{Drummer}), (d_{Guitarist}, Ringo_Starr), (John_Lennon, d_{Guitarist}), (John_Lennon, John_Lennon), (John_Lennon, d_{Bassist}), (John_Lennon, Paul_McCartney), (John_Lennon, d_{Drummer}), (John_Lennon, Ringo_Starr), (d_{Bassist}, d_{Guitarist}), (d_{Bassist}, John_Lennon), (d_{Bassist}, d_{Bassist}), (d_{Bassist}, Paul_McCartney), (d_{Bassist}, d_{Drummer}), (d_{Bassist}, Ringo_Starr), (Paul_McCartney, d_{Guitarist}), (Paul_McCartney, d_{Drummer}), (Paul_McCartney, Ringo_Starr), (d_{Drummer}, d_{Guitarist}), (d_{Drummer}, d_{Guitarist}), (d_{Drummer}, d_{Guitarist}), (d_{Drummer}, d_{Guitarist}), (d_{Drummer}, Ringo_Starr), (Ringo_Starr, d_{Guitarist}), (Ringo_Starr, d_{Drummer}), (Ringo_Starr, d_{Bassist}), (Ringo_Starr, d_{Drummer}), (Ringo_Starr, Ringo_Starr) \}
cetanswer(F(x, y), (\mathcal{T}, \mathcal{A})) = \{ (John_Lennon, John_Lennon), (Paul_McCartney, Paul_McCartney), (Paul_McCartney), (Paul_McCartney), (Ringo_Starr, Ringo_Starr, Ringo_Starr) \}
```

• $F = \exists x.managed_by(x, x)$

Because $(d_{Manager}, d_{Manager}) \in managed_by^{\mathcal{I}_{\mathcal{T}, \mathcal{A}}}$, $\mathcal{I}_{\mathcal{T}, \mathcal{A}}$ gives the answer "Yes".

But $(\mathcal{T},\mathcal{A})$ doesn't give the certain answer "Yes".

Question 4. Conjunctive queries over database and interpretation

(1)

The finite first-order interpretation \mathcal{I}_D corresponding to $\mathcal{D}.$

- $ID^{\mathcal{I}_{\mathcal{D}}} = \{2001, 2002, 2003, 2004, 30000160, 30000170, 30000180\}$
- Name^{ID} = { Jay_Chou, Jolin_Tsai, Stefanie_Sun, Ta-yu_Lo }
- StudentID ID = { 2001, 2002, 2003, 2004 }
- Since $\mathcal{I}_{\mathcal{D}} = \{2020, 2021, 2020\}$
- CourselD $^{\mathcal{I}_{\mathcal{D}}}$ = { 30000160, 30000170, 30000180 }
- Title $^{ID} = \{ KR\&P, PR\&CV, NLP \}$
- Person $^{\mathcal{I}_{\mathcal{D}}}$ = { (2001, Jay_Chou), (2002, Jolin_Tsai), (2003, Stefanie_Sun), (2004, Tayu_Lo) }

- Enrollment $\mathcal{I}_{\mathcal{D}} = \{ (2002, 2020), (2003, 2021), (2004, 2020) \}$
- Attendance $\mathcal{ID} = \{ (2001, 30000160), (2002, 30000160), (2002, 30000170), (2003, 30000180) \}$
- Course $\mathcal{ID} = \{ (30000160, KR\&P), (30000180, PR\&CV), (30000170, NLP) \}$

(2)

First-order queries f_Q :

- $F_a(x, y) = Person(x, y)$
- $F_b(x) = \exists y. \exists z. (Person(y, x) \land Attendance(y, z) \land Course(z, KR&P))$
- $F_c(x) = \exists y. \exists z. (Person(y, x) \land Enrollment(y, z) \land \forall c. \neg Attendance(y, c))$

 $F_a(x, y)$ and $F_b(x)$ are conjunctive queries but $F_c(x)$ is not conjunctive query, because there is a universal quantification $\forall c$ in $F_c(x)$.

(3)

Answer Q in the context of \mathcal{D} :

- answer(Q_a, D) = { (2001, Jay_Chou), (2002, Jolin_Tsai), (2003, Stefanie_Sun), (2004, Ta-yu_Lo) }
- answer(Q_b, D) = { Jay_Chou, Jolin_Tsai }
- answer(Q_c, D) = { Ta-yu_Lo }

Answer f_Q in the context of $I_{\mathcal{D}}$:

- answer(F_a , I_D) = { (2001, Jay_Chou), (2002, Jolin_Tsai), (2003, Stefanie_Sun), (2004, Ta-yu_Lo) }
- answer(F_b , I_D) = { Jay_Chou, Jolin_Tsai }
- answer(F_c , I_D) = { Ta-yu_Lo }

Question 5. Certain answers in different contexts

(1)

- cetanswer($r(x, y) \land Y(y), A) = \emptyset$
- cetanswer($\exists y(r(x, y) \land Y(y)), A) = \emptyset$
- cetanswer($\exists x, y(r(x, y) \land r(y, x)), A$) = "Yes"
- cetanswer($\exists z, w(r(x, y) \land r(y, z) \land r(z, x) \land r(z, w) \land W(w)), A) = \emptyset$

- cetanswer($r(x, y) \land Y(y), A$) = { (Jay_Chou, Jolin_Tsai), (Jolin_Tsai, Stefanie_Sun), (Stefanie_Sun, Jay_Chou), (Jolin_Tsai, Jolin_Tsai), (Stefanie_Sun, Stefanie_Sun) }
- cetanswer($\exists y(r(x, y) \land Y(y)), A) = \{ Jay_Chou, Jolin_Tsai, Stefanie_Sun \}$
- cetanswer($\exists x, y(r(x, y) \land r(y, x)), A$) = "Yes"
- cetanswer($\exists z, w(r(x, y) \land r(y, z) \land r(z, x) \land r(z, w) \land W(w)), A) = { (Jay_Chou, Jolin_Tsai), (Stefanie_Sun, Jay_Chou), (Jolin_Tsai, Jolin_Tsai), (Jolin_Tsai, Stefanie_Sun) }$

Question 6 (with 1 bonus mark). Simpleness of ABox

This doesn't affect the lower bound of the data complexity results.

Lower bound:

• \mathcal{ALC} : coNP-hard

• \mathcal{EL} : P-hard

• DL-Lite: AC⁰

Because questions like non-3-colorability and path system accessibility can be still reduced into \mathcal{ALC} and \mathcal{EL} CQ-entailment problem by the way same with simple ABoxes.

This doesn't affect the upper bound of the data complexity results, too.

Upper bound:

- \mathcal{ALC} : coNP-complete, because we can still use tableau algorithm to solve the problem.
- \mathcal{EL} : P-complete, because we can reduce it to Datalog query entailment with PTime data complexity.
- $\mathrm{DL\text{-}Lite}$: AC^0 , because we can reduce it to entailment of their FO-rewriting $q_{\mathcal{T}}$.

So this doesn't affect the data complexity result.

Question 7 (with 1 bonus mark). k-colorability

It is possible to show that the problem of conjunctive query entailment (CQ-entailment) in ALC is coNP-hard w.r.t. data complexity using a reduction from non-k-colorability in graphs.

We can consider the following \mathcal{ALC} TBox and Boolean CQ:

$$\mathcal{T} = \{\top \sqsubseteq \bigsqcup_{C \in \operatorname{colors}} C, C \sqcap \exists r. C \sqsubseteq D \text{ for } C \in \operatorname{colors} \}$$

$$q = \exists x. D(x)$$

And we translated the input graph G=(V,E) into the ABox:

$$\mathcal{A} = \{(u,v): r | \{u,v\} \in E\}$$

Then it is straightforward to prove that $(\mathcal{T}, \mathcal{A}) \models q$ if and only if \mathcal{A} is not k-colourable.

Because 3-colorability is NP-hard, thus k-colorability is NP-hard also. So we can know that the query entailment problem for conjunctive queries is coNP-hard w.r.t data complexity in \mathcal{ALC} .

But if we let k is fixed, it depends on the value of k whether the query entailment problem for conjunctive queries is coNP-hard w.r.t data complexity in \mathcal{ALC} .

For example, let k=1 and then a graph is 1-colorable if and only if it is totally disconnected, that is all its vertices are isolated. It can be identified in PTime. So it can not prove the proposition.

Let k=2 and then a graph is 2-colorable if and only if its vertices having same color can be taken as disjoint sets. Thus, A graph is 2-colorable if and only if it is 2-colorable, which can be also solved in PTime. So it can not prove the proposition.

Let k=3 and then 3-colorability is NP-hard and it can prove that the query entailment problem for conjunctive queries is coNP-hard w.r.t data complexity in \mathcal{ALC} .

So if k is fixed, it depends on the value of k whether the query entailment problem for conjunctive queries is coNP-hard w.r.t data complexity in \mathcal{ALC} .