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) ∃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.⊤(Jay_Chou): Yes
- (i) ∃hasFriend.∃dancesWith.Song(Jay_Chou): Yes
- (j) ∃hasFriend.Composer(Jay_Chou): No
- (k) ∃hasFriend.{Jay_Chou}(Vincent_Fang): No
- (l) 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.⊤(Jay_Chou): Yes
- (i) ∃hasFriend.∃dancesWith.Song(Jay_Chou): Yes
- (j) ∃hasFriend.Composer(Jay_Chou): Don't know
- (k) ∃hasFriend.{Jay_Chou}(Vincent_Fang): Don't know
- (l) 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}) = ∅
- (c) cetanswer(F3(x, y), D_{music}) = { (Jay_Chou, Vincent_Fang), (Jay_Chou, Will_Liu) }
- (d) cetanswer(F4(x), D_{music}) = ∅

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_{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) ∃hasFriend.∃dancesWith.Song(Jay_Chou): Yes
- (j) ∃hasFriend.Composer(Jay_Chou): Don't know
- (k) ∃hasFriend.{Jay_Chou}(Vincent_Fang): Don't know
- (l) 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

- $S(d_{Guitarist})$ = { 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) = ∅

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:

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

(2)

- ∃plays_for.RockBand(John_Lennon):

Because (John_Lennon, $d_{RockBand}$) ∈ 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 ∈ Manager$^{\mathcal{I}_{\mathcal{T},\mathcal{A}}}$ and (Paul_McCartney, x) ∈ managed_by$^{\mathcal{I}_{\mathcal{T},\mathcal{A}}}$ 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".

- ∃plays_for.∃managed_by.Manager(Ringo_Starr)

Because (Ringo_Starr, $d_{RockBand}$) ∈ plays_for$^{\mathcal{I}_{\mathcal{T},\mathcal{A}}}$ and ($d_{RockBand}$, $d_{Manager}$) ∈ 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".

(3)

- F(x, y) = ∃z.(plays_for(x, z) ∧ 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, John_Lennon), (Paul_McCartney, $d_{Bassist}$), (Paul_McCartney, Paul_McCartney), (Paul_McCartney, $d_{Drummer}$), (Paul_McCartney, Ringo_Starr), ($d_{Drummer}$, $d_{Guitarist}$), ($d_{Drummer}$, John_Lennon), ($d_{Drummer}$, $d_{Bassist}$), ($d_{Drummer}$, Paul_McCartney), ($d_{Drummer}$, $d_{Drummer}$), ($d_{Drummer}$, Ringo_Starr), (Ringo_Starr, $d_{Guitarist}$), (Ringo_Starr, John_Lennon), (Ringo_Starr, $d_{Bassist}$), (Ringo_Starr, Paul_McCartney), (Ringo_Starr, $d_{Drummer}$), (Ringo_Starr, Ringo_Starr) }

cetanswer(F(x, y), ($\mathcal{T}, \mathcal{A}$)) = { (John_Lennon, John_Lennon), (Paul_McCartney, Paul_McCartney), (Ringo_Starr, Ringo_Starr) }

- F = ∃x.managed_by(x, x)

Because ($d_{Manager}$, $d_{Manager}$) ∈ 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 $\mathcal{I}_{\mathcal{D}}$ corresponding to $\mathcal{D}$.

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

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

(2)

First-order queries f_Q:

- F_a(x, y) = Person(x, y)
- F_b(x) = ∃y.∃z.(Person(y, x) ∧ Attendance(y, z) ∧ Course(z, KR&P))
- F_c(x) = ∃y.∃z.(Person(y, x) ∧ Enrollment(y, z) ∧ ∀c.¬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 ∀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, $I_\mathcal{D}$) = { (2001, Jay_Chou), (2002, Jolin_Tsai), (2003, Stefanie_Sun), (2004, Ta-yu_Lo) }
- answer(F_b, $I_\mathcal{D}$) = { Jay_Chou, Jolin_Tsai }
- answer(F_c, $I_\mathcal{D}$) = { Ta-yu_Lo }

Question 5. Certain answers in different contexts

(1)

- cetanswer(r(x, y) ∧ Y (y), $\mathcal{A}$) = ∅
- cetanswer(∃y(r(x, y) ∧ Y (y)), $\mathcal{A}$) = ∅
- cetanswer(∃x, y(r(x, y) ∧ r(y, x)), $\mathcal{A}$) = "Yes"
- cetanswer(∃z, w(r(x, y) ∧ r(y, z) ∧ r(z, x) ∧ r(z, w) ∧ W(w)), $\mathcal{A}$) = ∅

(2)

- cetanswer(r(x, y) ∧ Y (y), $\mathcal{A}$) = { (Jay_Chou, Jolin_Tsai), (Jolin_Tsai, Stefanie_Sun), (Stefanie_Sun, Jay_Chou), (Jolin_Tsai, Jolin_Tsai), (Stefanie_Sun, Stefanie_Sun) }
- cetanswer(∃y(r(x, y) ∧ Y (y)), $\mathcal{A}$) = { Jay_Chou, Jolin_Tsai, Stefanie_Sun }
- cetanswer(∃x, y(r(x, y) ∧ r(y, x)), $\mathcal{A}$) = "Yes"
- cetanswer(∃z, w(r(x, y) ∧ r(y, z) ∧ r(z, x) ∧ r(z, w) ∧ W(w)), $\mathcal{A}$) = { (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
- $\mathrm{DL\text{-}Lite}$: AC^0

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:

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

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

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 a bipartite graph, 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 $\mathcal{ALC}$.