Complexity Theory Exam (VU 181.142)

21 January, 2022

Variants of Subsumption

Definition 1. [terms, atoms, literals, clauses]

- A *term* t is either a variable (denoted by u, v, w, x, y, z, etc.) or a constant (denoted by a, b, c, etc.).
- An *atom* A is an expression of the form $P(t_1, \ldots, t_k)$, where P is a predicate symbol of arity k and $t_1, \ldots, t_k$ are terms.
- A *literal* L is either an atom A or a negated atom $\neg A$.
- A *clause* C is a disjunction of literals. For simplicity, we may denote clauses as *sets of literals*, i.e., we write $C = \{L_1, \ldots, L_n\}$ rather than $C = L_1 \vee \cdots \vee L_n$.

Definition 2. [substitution]

- A *substitution* λ is a mapping which maps variables to terms. We write

$$\lambda = \{x_1 \leftarrow t_1, \ldots, x_m \leftarrow t_m\}$$

  to denote the substitution which maps x_i to t_i for each $i \in \{1, \ldots, m\}$.
- We denote the application of a substitution λ to an expression E (where E may be a term, an atom, a literal, or a clause) as $E\lambda$, i.e.: we simultaneously replace every occurrence of each variable x_i in E by t_i.

Definition 3. [subsumption]
Let $C = \{L_1, \ldots, L_k\}$ and $D = \{M_1, \ldots, M_\ell\}$ be clauses. We say that "C subsumes D" if the following condition holds: there exist a substitution λ on the variables occurring in C which simultaneously maps every literal L_i in C to some literal M_j in D, i.e., $\{L_1\lambda, \ldots, L_k\lambda\} \subseteq \{M_1, \ldots, M_\ell\}$, i.e., $C\lambda \subseteq D$, for short.

Example 1. [clauses, substitutions, subsumption] Consider the predicate symbol R with arity 3 and S with arity 2, and define the clauses C and D as follows:

$$C = \{R(x_1, x_2, x_3), R(x_1, x_3, x_4), S(x_4, x_2), S(x_3, x_2), S(x_2, x_1), S(x_1, x_2)\}$$
$$D = \{R(b_1, b_1, b_2), R(b_1, b_2, b_2), S(b_2, b_1), S(b_1, b_1), S(b_1, b_3)\}.$$

We claim that C subsumes D. Indeed, consider the substitution $\lambda = \{x_1 \leftarrow b_1,\ x_2 \leftarrow b_1,\ x_3 \leftarrow b_2,\ x_4 \leftarrow b_2\}$. It can be easily checked that $C\lambda \subseteq D$ holds.

Definition 4. [ground substitution, ground instance]

- Let E be an expression (where E may be a term, an atom, a literal, or a clause) and let $H = \{b_1, \ldots, b_m\}$ be a universe. We say that λ is a *ground substitution* (over H) for E if λ maps every variable in E to some element $b_i \in H$.
- Let E and E' be expressions. We say that E' is a *ground instance* of E (over H) if there exists a ground substitution λ (over H), s.t. $E' = E\lambda$.

Definition 5. [H-subsumption]
Let $C = \{L_1, \ldots, L_k\}$ and $D = \{M_1, \ldots, M_\ell\}$ be clauses and let $H = \{b_1, \ldots, b_m\}$ denote a universe. We say that "C H-subsumes D" if the following condition holds: for every ground substitution μ (over H) for D, there exists a ground substitution λ (over H) for C, s.t. $C\lambda \subseteq D\mu$. In other words, C subsumes every ground instance of D over H.

Definition 6. [H-SUBSUMPTION problem]
INSTANCE: Two clauses $C = \{L_1, \ldots, L_k\}$ and $D = \{M_1, \ldots, M_\ell\}$ and a universe $H = \{b_1, \ldots, b_m\}$.
QUESTION: Does C H-subsume D? That is, does C subsume every ground instance of D over H?

Theorem. *The H-SUBSUMPTION problem is $\Pi_2 P$-complete.*

Problem reduction for the $\Pi_2 P$-hardness proof. We prove the $\Sigma_2 P$-hardness of the co-problem of H-SUBSUMPTION. Let an arbitrary instance of the QSAT_2 problem be given by the formula $\psi = \exists X \forall Y \varphi(X, Y)$ with $X = \{x_1, \ldots, x_k\}$ and $Y = \{y_1, \ldots, y_\ell\}$, where $\varphi(X, Y)$ is in 3-DNF, i.e.: $\varphi(X, Y) = \varphi_1 \vee \cdots \vee \varphi_n$ with $\varphi_i = l_{i1} \wedge l_{i2} \wedge l_{i3}$ for $i \in \{1, \ldots, n\}$, s.t. each $l_{\alpha\beta}$ is a literal over the propositional variables in $X \cup Y$. Then we define an instance of the co-H-SUBSUMPTION problem as follows:

- Let P, Q, R be predicate symbols, s.t. P has arity k, Q has arity 2, and R has arity 3. Moreover, we consider the universe $H = \{a_0, a_1\}$.
- We define the clause C with variables $\{u_1, \ldots, u_k, u'_1, \ldots, u'_k, v_1, \ldots, v_\ell, v'_1, \ldots, v'_\ell\}$ and the clause D with variables $\{w_1, \ldots, w_k\}$ as follows:
- $C = \{P(u_1, \ldots, u_k)\} \cup$
  $\{Q(u_1, u'_1), \ldots, Q(u_k, u'_k), Q(v_1, v'_1), \ldots, Q(v_\ell, v'_\ell)\} \cup$
  $\{R(l^*_{11}, l^*_{12}, l^*_{13}), \ldots, R(l^*_{n1}, l^*_{n2}, l^*_{n3})\}$, where $l^*_{\alpha\beta}$ is defined as

  $l^*_{\alpha\beta} = u_\gamma$ if $l_{\alpha\beta} = x_\gamma$, $l^*_{\alpha\beta} = u'_\gamma$ if $l_{\alpha\beta} = \neg x_\gamma$,
  $l^*_{\alpha\beta} = v_\gamma$ if $l_{\alpha\beta} = y_\gamma$, $l^*_{\alpha\beta} = v'_\gamma$ if $l_{\alpha\beta} = \neg y_\gamma$.
- $D = \{P(w_1, \ldots, w_k)\} \cup \{Q(a_0, a_1), Q(a_1, a_0)\} \cup$
  $\{R(a_0, a_0, a_0), R(a_0, a_0, a_1), R(a_0, a_1, a_0), R(a_0, a_1, a_1)\} \cup$
  $\{R(a_1, a_0, a_0), R(a_1, a_0, a_1), R(a_1, a_1, a_0)\}$.

Example 2. [subsumption vs. H-subsumption] Let $C = \{P(x_1, x_2), Q(x_1), Q(x_2)\}$ and $D = \{P(y_1, y_2), Q(a), Q(b)\}$. Then the following relationships hold:

- C does not subsume D, since there is no substitution λ on the variables $\{x_1, x_2\}$ in C, s.t., on the one hand $P(x_1, x_2)\lambda \in D$ and, on the other hand, also $Q(x_1)\lambda \in D$ and $Q(x_2)\lambda \in D$ hold.

- For $H = \{a, b\}$, C H-subsumes D. To see this, consider an arbitrary ground instance $D\mu$ of D over H, i.e.: $y_1\mu \in \{a, b\}$ and $y_2\mu \in \{a, b\}$. Then C subsumes $D\mu$ since $C\lambda \subseteq D\mu$ clearly holds for $\lambda = \{x_1 \leftarrow y_1\mu, x_2 \leftarrow y_2\mu\}$, i.e., on the one hand, $P(x_1, x_2)\lambda = P(y_1, y_2)\mu$ and, on the other hand, $Q(x_1)\lambda \in \{Q(a), Q(b)\}$ and $Q(x_2)\lambda \in \{Q(a), Q(b)\}$ hold.

- For $H = \{a, b, c\}$, C does not H-subsume D. To see this, consider the ground substitution $\mu = \{y_1 \leftarrow c, y_2 \leftarrow c\}$ over H. Then the ground instance $D\mu$ of D is clearly not subsumed by C.

Example 3. [Problem reduction] Consider the QSAT$_2$-formula $\psi = \exists X \forall Y \varphi(X, Y)$ with $X = \{x_1, x_2\}$, $Y = \{y_1, y_2\}$, and $\varphi(X, Y) = (x_1 \wedge y_1 \wedge \neg x_2) \vee (x_1 \wedge y_2 \wedge x_2)$. The above problem reduction yields the following clauses C and D over the universe $H = \{a_0, a_1\}$:

- $C = \{P(u_1, u_2)\} \cup$
  $\{Q(u_1, u_1'), Q(u_2, u_2'), Q(v_1, v_1'), Q(v_2, v_2')\} \cup$
  $\{R(u_1, v_1, u_2'), R(u_1, v_2, u_2)\}$,
- $D = \{P(w_1, w_2)\} \cup \{Q(a_0, a_1), Q(a_1, a_0)\} \cup$
  $\{R(a_0, a_0, a_0), R(a_0, a_0, a_1), R(a_0, a_1, a_0), R(a_0, a_1, a_1)\} \cup$
  $\{R(a_1, a_0, a_0), R(a_1, a_0, a_1), R(a_1, a_1, a_0)\}$.

In this example, the formula ψ is not true, i.e., no matter what truth value we assign to the variables x_1 and x_2, there always exists a truth value assignment to y_1 and y_2, s.t. none of the disjuncts in the 3-DNF φ is true. On the other hand, C does H-subsume D, i.e., no matter what substitution μ on the variables in D we choose, there always exists a substitution λ on the variables in C with $C\lambda \subseteq D\mu$.

For instance, consider the truth assignment $I(x_1) = \text{true}$ and $I(x_2) = \text{false}$. There exists an extension J of I to $\{y_1, y_2\}$ which falsifies φ, namely $J(y_1) = \text{false}$ and $J(y_2) = \text{true}$. Now consider the substitution μ "corresponding to I", i.e., $\mu = \{w_1 \leftarrow a_1, w_2 \leftarrow a_0\}$. We claim that $D\mu$ is subsumed by C. Indeed, consider the substitution λ "corresponding to J", i.e., $\lambda = \{u_1 \leftarrow a_1, u_1' \leftarrow a_0, u_2 \leftarrow a_0, u_2' \leftarrow a_1, v_1 \leftarrow a_0, v_1' \leftarrow a_1, v_2 \leftarrow a_1, v_2' \leftarrow a_0\}$. Obviously, $C\lambda \subseteq D\mu$ holds.

Some Hints and Remarks.

- In the problem reduction, the first-order variables u_i, u_i', v_i, v_i' "encode" the truth value of the propositional literals x_i, $\neg x_i$, y_i, $\neg y_i$. The first-order variables w_i, w_i' also "encode" the truth value of x_i, $\neg x_i$. For instance, we have the following correspondence between an assignment J on ψ and a substitution λ (with $C\lambda \subseteq D\mu$):

  $J(x_i) = \text{true}$ iff $u_i\lambda = a_1$ iff $u_i'\lambda = a_0$ and

  $J(y_i) = \text{true}$ iff $v_i\lambda = a_1$ iff $v_i'\lambda = a_0$

  (analogous equivalences hold if J assigns the truth value false to x_i or y_i).

- For the equivalence between the instance of QSAT$_2$ and the instance of co-H-SUBSUMPTION it is important to note that φ is false in some assignment J iff the three literals in each disjunct of the 3-DNF are mapped to a value combination different from (true, true, true). On the other hand, $C\lambda$ subsumes $D\mu$ iff

  (1) $P(u_1, \ldots, u_k)\lambda = P(w_1, \ldots, w_k)\mu$,

  (2) the variables u_i and u_i' (likewise v_i and v_i') are mapped to distinct values, and

  (3) none of the atoms $R(l_{i1}^*, l_{i2}^*, l_{i3}^*)\lambda$ has the value $R(a_1, a_1, a_1)$.

- Please keep in mind that all the above comments are *explanations* of the intuition of the problem reduction. They are *not proofs*. When you are requested to prove the correctness of the problem reduction, you are not allowed to refer to these explanations. *Your proofs have to be self-contained*! That is, for any property that you use in your proof, make it perfectly clear why this property holds (e.g., "by the problem reduction", "by assumption X", "by definition X", etc.)

Name: ....

Matr-Nr: ....

Question 1 (15 credits) Prove the first direction of the correctness of the above problem reduction, namely:

If there exists a truth assignment I on the variables in X, s.t. for all extensions J of I to the variables in Y, $\varphi(X,Y)$ is true in J, then C *does not H-subsume* D.

Solution to Question 1.

Name: ....

Matr-Nr: ....

Question 2 (15 credits) Prove the second direction of the correctness of the above problem reduction, namely:

If C *does not H-subsume* D, then there exists a truth assignment I on the variables in X, s.t. for all extensions J of I to the variables in Y, $\varphi(X, Y)$ is true in J.

Solution to Question 2.