Estructuras de Herbrand, Unificación y Resolución

Pablo Barceló

Estructuras genéricas

Dado conjunto de oraciones Σ para el cual deseamos verificar satisfacibilidad.

- Quizás no debemos inspeccionar todo el espacio de estructuras.
- Podríamos dejar afuera interpretaciones exóticas (números naturales, reales, u otras estructuras matemáticas) que no tienen que ver con el dominio de discurso.
- Veremos que en algunos casos es posible restringirse a una clase de estructuras genérica, que está muy ligada a la sintaxis (vocabulario) de Σ.

Esta idea fue desarrollada a prinicipos de los años 30 por Jacques Herbrand (¿Qué dice Wikipedia acerca de él?).

Universo de Herbrand

Sea $\mathcal{L}$ un vocabulario con al menos una constante.

El universo de Herbrand de $\mathcal{L}$, denotado por $U_{\mathcal{L}}$, es el conjunto de todos los términos sin variables libres de $\mathcal{L}$.

Ejemplo: Sea $\mathcal{L}$ un vocabulario que contiene a la constante a, la función binaria g, y la relación R, entonces:

$$U_{\mathcal{L}} = \{a, g(a, a), g(a, g(a, a)), g(g(a, a), a), \ldots\}.$$

Si $\mathcal{L}$ no tiene constantes, simplemente agregamos una.

Observación: $U_{\mathcal{L}}$ es infinito si y sólo si $\mathcal{L}$ contiene al menos un símbolo de función.

Base de Herbrand

La base de Herbrand $B_{\mathcal{L}}$ asociada al vocabulario $\mathcal{L}$ es el conjunto de todas las oraciones atómicas que se pueden formar con los elementos de $U_{\mathcal{L}}$.

La base de Herbrand asociada a la relación n-aria R es el conjunto de todas las n-tuplas $(a_1, \ldots, a_n) \in (U_{\mathcal{L}})^n$ (note que para cada una de éstas se tiene que $R(a_1, \ldots, a_n) \in B_{\mathcal{L}}$).

Ejemplo: Sea $\mathcal{L}$ un vocabulario con dos constantes c, d, una función unaria f, y una relación binaria R.

Entonces $U_{\mathcal{L}} = \{c, d, f(c), f(d), f(f(c)), f(f(d)), \ldots\}$, y la base de Herbrand es

$$\begin{aligned} B_{\mathcal{L}} \ = \ \{ & R(c,c), R(c,d), R(d,d).R(d,c), \\ & R(f(c),c), R(f(c),d), R(c,f(c)), R(d,f(c)), \ldots\}. \end{aligned}$$

Estructura de Herbrand

Finalmente llegamos a nuestro concepto más importante:

Una $\mathcal{L}$-estructura de Herbrand es una estructura tal que:

- Su dominio es $U_{\mathcal{L}}$.
- Cada constante c en $\mathcal{L}$ se interpreta como c.
- Cada función m-aria f asigna a la tupla $(t_1, \ldots, t_m) \in (U_{\mathcal{L}})^m$ el elemento $f(t_1, \ldots, t_m)$.
- Cada relación n-aria R se interpreta como un subconjunto de la base de Herbrand asociada a R.

Note que todas las $\mathcal{L}$-estructuras de Herbrand comparten el dominio y la interpretación de constantes y funciones.

Notación: Si existe estructura de Herbrand $\mathcal{A}$ que satisface Σ, decimos que Σ tiene un modelo de Herbrand.

Oraciones universales y estructuras de Herbrand

Pregunta: ¿Existe alguna clase (interesante) de oraciones para los cuales verificar satisfacibilidad se reduce a verificar satisfacibilidad sólo con respecto a modelos de Herbrand?

Oraciones universales y estructuras de Herbrand

Pregunta: ¿Existe alguna clase (interesante) de oraciones para los cuales verificar satisfacibilidad se reduce a verificar satisfacibilidad sólo con respecto a modelos de Herbrand?

La respuesta afirmativa fue dada también por Herbrand: Una oración es universal si es de la forma $\forall x_1 \ldots \forall x_n \psi$, donde ψ no tiene cuantificación (asumimos que ψ está en CNF, i.e. ψ es una conjunción de cláusulas).

Teorema (Herbrand)

Sea Σ un conjunto de oraciones universales que no mencionan el símbolo de igualdad $=$. Entonces Σ es insatisfacible si y sólo si Σ no tiene un modelo de Herbrand.

Veremos que las oraciones universales forman clase *interesante*.

Demostración del teorema de Herbrand

Una dirección del teorema es trivial: Si Σ es satisfacible por una estructura de Herbrand entonces es satisfacible.

Pregunta: ¿Cómo podemos resolver la otra dirección?

Demostración del teorema de Herbrand

Una dirección del teorema es trivial: Si Σ es satisfacible por una estructura de Herbrand entonces es satisfacible.

Pregunta: ¿Cómo podemos resolver la otra dirección?

Asuma Σ es satisfacible por $\mathcal{L}$-estructura $\mathcal{A}$.

Construya $\mathcal{L}$-estructura de Herbrand $\mathcal{A}_H$ tal que para cada relación n-aria R, la interpretación de R en $\mathcal{A}_H$ es el conjunto:

$$\{(t_1, \ldots, t_n) \in (U_{\mathcal{L}})^n \mid \mathcal{A} \models R(t_1, \ldots, t_n)\}.$$

Ejercicio: Demuestre que $\mathcal{A}_H \models \Sigma$.

Teorema de Herbrand y otro tipo de oraciones

Pregunta: ¿Qué sucede si en el teorema de Herbrand sacamos la condición de que las oraciones sean universales?

Teorema de Herbrand y otro tipo de oraciones

Pregunta: ¿Qué sucede si en el teorema de Herbrand sacamos la condición de que las oraciones sean universales?

Entonces el teorema no es cierto.

Considere el conjunto Σ de oraciones $\{P(a), \exists x \neg P(x)\}$.

Es claro que Σ es satisfacible.

Pero las únicas estructuras de Herbrand posibles son las que interpretan a P como $\emptyset$ y $\{a\}$, respectivamente. Ninguna de éstas satisface a Σ.

Teorema de Herbrand y otro tipo de oraciones

Pregunta: ¿Qué sucede si en el teorema de Herbrand sacamos la condición de que las oraciones sean universales?

Entonces el teorema no es cierto.

Considere el conjunto Σ de oraciones $\{P(a), \exists x \neg P(x)\}$.

Es claro que Σ es satisfacible.

Pero las únicas estructuras de Herbrand posibles son las que interpretan a P como $\emptyset$ y $\{a\}$, respectivamente. Ninguna de éstas satisface a Σ.

Más adelante veremos cómo resolver este inconveniente.

Teorema de Herbrand y otro tipo de oraciones

Pregunta: ¿Qué sucede si en el teorema de Herbrand sacamos la condición de que las oraciones no mencionen igualdad?

Teorema de Herbrand y otro tipo de oraciones

Pregunta: ¿Qué sucede si en el teorema de Herbrand sacamos la condición de que las oraciones no mencionen igualdad?

Entonces el teorema no es cierto.

Considere la oración $(c = f(c))$.

Es claro que Σ es satisfacible.

Pero en toda estructura de Herbrand $c \neq f(c)$.

Pregunta: ¿Qué sucede si en el teorema de Herbrand sacamos la condición de que las oraciones no mencionen igualdad?

Entonces el teorema no es cierto.

Considere la oración $(c = f(c))$.

Es claro que Σ es satisfacible.

Pero en toda estructura de Herbrand $c \neq f(c)$.

Más adelante veremos cómo este inconveniente puede ser también resuelto.

De ahora en adelante, una fórmula universal $\forall \bar{x} \phi(\bar{x})$ se representará por la conjunción de cláusulas $\phi(\bar{x})$, y asumiremos implícitamente la cuantificación universal.

Una instanciación de la cláusula $\alpha(\bar{x})$ es una oración obtenida desde $\alpha(\bar{x})$ reemplazando cada variable libre x por un término t en $U_{\mathcal{L}}$.

Ejemplo: Sea $U_{\mathcal{L}} = \{c, d, f(c), f(d), f(f(c)), f(f(d)), \dots\}$ y cláusula $\{P(x), R(c, y)\}$. Una posible instanciación es

$$\{P(a), R(c, f(f(f(d))))\}.$$

Teorema de Herbrand y resolución proposicional

Para conjunto de cláusulas Σ sin igualdad sea Σ' el conjunto de todas las instanciaciones de las cláusulas en Σ.

Por el teorema de Herbrand, Σ es satisfacible si y sólo si Σ' tiene un modelo de Herbrand.

Además, Σ' es *equivalente* a un conjunto (posiblemente infinito) de fórmulas proposicionales.

Idea: Ocupar resolución proposicional para comprobar la satisfacibilidad de Σ'.

Teorema de Herbrand y resolución proposicional

Para conjunto de cláusulas Σ sin igualdad sea Σ' el conjunto de todas las instanciaciones de las cláusulas en Σ.

Por el teorema de Herbrand, Σ es satisfacible si y sólo si Σ' tiene un modelo de Herbrand.

Además, Σ' es *equivalente* a un conjunto (posiblemente infinito) de fórmulas proposicionales.

Idea: Ocupar resolución proposicional para comprobar la satisfacibilidad de Σ'.

Salvo por un pequeño detalle (cuál?).

La idea anterior no puede ser utilizada directamente.

Podemos tratar de refinar este procedimiento manteniendo nuestra búsqueda de refutaciones lo más *general* posible.

Idea: Sean cláusulas $\{P(x,z), \neg Q(x)\}$ y $\{\neg P(b,y), R(b,f(y))\}$.

Podemos hacer resolución proposicional a partir de cualquier valor asignado a las variables.

- Si $x = b$ e $y = z$ obtenemos $\{\neg Q(b), R(b,f(y))\}$.
- Si $x = b$ e $y = z = f(f(a))$ obtenemos $\{\neg Q(b), R(b,f(f(f(a))))\}$.

Intentaremos formalizar esta idea.

Unificación: Sustituciones

Una sustitución θ es un conjunto *finito* de la forma

$$\{x_1/t_1, \ldots, x_n/t_n\},$$

tal que cada para cada $i \leq n$, x_i es una variable y t_i es un término distinto de x_i.

Observaciones:

- Cada elemento x_i/t_i se llama ligazón.
- La sustitución es libre de variables si cada t_i es un término sin variables.
- La sustitución es variable por variable si cada t_i es una variable.

Unificación: Sustituciones

Una expresión es un término, una fórmula atómica o una cláusula.

Dada expresión E y sustitución $\theta = \{x_1/t_1, \ldots, x_n/t_n\}$:

Denotamos por $E\theta$ a la expresión obtenida desde E al reemplazar *simultáneamente* cada ocurrencia de x_i en E por t_i.

Unificación: Sustituciones

Una expresión es un término, una fórmula atómica o una cláusula.

Dada expresión E y sustitución $\theta = \{x_1/t_1, \ldots, x_n/t_n\}$:

Denotamos por $E\theta$ a la expresión obtenida desde E al reemplazar *simultáneamente* cada ocurrencia de x_i en E por t_i.

Ejemplo: Sea $\theta = \{x/(1+y), y/0, z/u\}$ y $E = (1 \cdot (x + y + 0) + z)$.

Unificación: Sustituciones

Una expresión es un término, una fórmula atómica o una cláusula.

Dada expresión E y sustitución $\theta = \{x_1/t_1, \ldots, x_n/t_n\}$:

Denotamos por $E\theta$ a la expresión obtenida desde E al reemplazar *simultáneamente* cada ocurrencia de x_i en E por t_i.

Ejemplo: Sea $\theta = \{x/(1+y), y/0, z/u\}$ y $E = (1 \cdot (x+y+0)+z)$.

Entonces $E\theta = (1 \cdot ((1+y)+0+0)+u)$.

Unificación: Sustituciones

Una expresión es un término, una fórmula atómica o una cláusula.

Dada expresión E y sustitución $\theta = \{x_1/t_1, \ldots, x_n/t_n\}$:

Denotamos por $E\theta$ a la expresión obtenida desde E al reemplazar *simultáneamente* cada ocurrencia de x_i en E por t_i.

Ejemplo: Sea $\theta = \{x/(1+y), y/0, z/u\}$ y $E = (1 \cdot (x+y+0)+z)$.

Entonces $E\theta = (1 \cdot ((1+y)+0+0)+u)$.

Ejemplo: Sea $\theta = \{x/b, y/x\}$ y $E = P(x, y, f(a))$.

Unificación: Sustituciones

Una expresión es un término, una fórmula atómica o una cláusula.

Dada expresión E y sustitución $\theta = \{x_1/t_1, \dots, x_n/t_n\}$:

Denotamos por $E\theta$ a la expresión obtenida desde E al reemplazar *simultáneamente* cada ocurrencia de x_i en E por t_i.

Ejemplo: Sea $\theta = \{x/(1+y), y/0, z/u\}$ y $E = (1 \cdot (x+y+0)+z)$.

Entonces $E\theta = (1 \cdot ((1+y)+0+0)+u)$.

Ejemplo: Sea $\theta = \{x/b, y/x\}$ y $E = P(x, y, f(a))$.

Entonces $E\theta = P(b, x, f(a))$.

Regla de resolución de primer orden

Dados literales l y p de la forma $P(\ldots)$, y cláusulas $C_1 \cup \{l\}$ y $C_2 \cup \{\neg p\}$.

Asumimos que los conjuntos de variables mencionados en cada cláusula son disjuntos (si no, sólo renombramos variables).

Regla de resolución: Si existe sustitución θ tal que $l\theta = p\theta$ entonces:

$$\frac{\begin{array}{c} C_1 \cup \{l\} \\ C_2 \cup \{\neg p\} \end{array}}{(C_1 \cup C_2)\theta}$$

La regla es correcta: $C_1 \cup \{l\},\ C_2 \cup \{\neg p\} \models (C_1 \cup C_2)\theta$.

Ejemplo de demostración por resolución

Ejemplo: Considere Σ conjunto de cláusulas:

$$Sobre(a, b), \quad Sobre(b, c), \quad Azul(a), \quad \neg Azul(c).$$

Queremos saber si la consulta

$$\exists x \exists y (Sobre(x, y) \wedge Azul(x) \wedge \neg Azul(y))$$

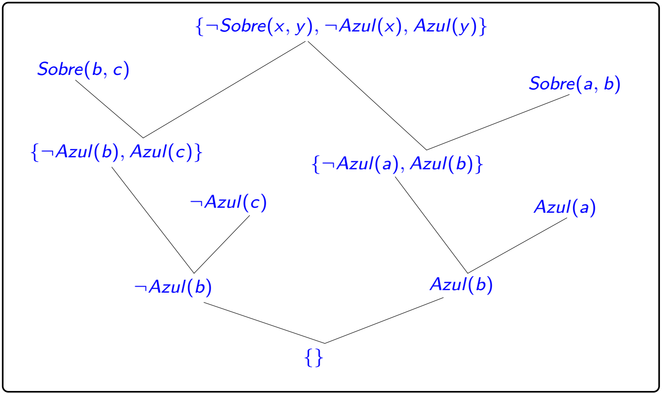
es consecuencia lógica de Σ.

Basta demostrar que $\Sigma \cup \{\neg Sobre(x, y), \neg Azul(x), Azul(y)\}$ es insatisfacible.

Usaremos resolución.

Un ejemplo de resolución

Ejemplo: (Continuación).



Obtención de respuestas

Sea Q una consulta existencial con variables libres.

Quisiéramos que el proceso de resolución nos diera también los *valores* que pueden tomar las variables libres para que Q sea consecuencia lógica de la base de conocimiento.

Es decir, queremos obtener el conjunto $\{\bar{a} \mid \Sigma \models Q(\bar{a})\}$.

Observación: El ejemplo anterior demuestra que $\exists x \phi(x)$ puede ser implicado por la base de conocimiento, pero $\phi(t)$ no lo es para ningún t.

Obtención de respuestas

La estrategia que se utiliza es reemplazar la consulta (existencial) $\phi(\bar{x})$ con $\phi(\bar{x}) \wedge \neg Ans(\bar{x})$, donde $Ans(\bar{x})$ es un predicado nuevo.

Es decir, realizamos resolución desde $\Sigma \cup \{\phi(\bar{x}) \rightarrow Ans(\bar{x})\}$.

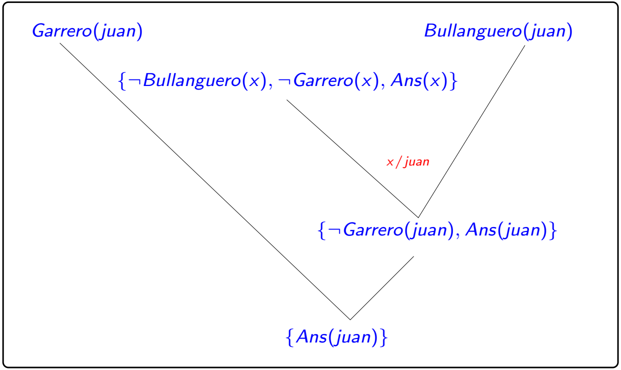
Dado que *Ans* no aparece en otras cláusulas de Σ, no será ahora siempre posible obtener la cláusula vacía desde la consulta modificada.

Lo que se hace es terminar la demostración por resolución una vez que se obtiene una cláusula que sólo contiene *Ans*.

Obtención de respuestas: Ejemplo

Ejemplo: Asuma que Σ contiene los átomos $Garrero(juan)$ y $Bullanguero(juan)$. La consulta es: $Garrero(x) \wedge Bullanguero(x)$.

Realizamos el siguiente procedimiento para obtener respuestas:



Obtención de respuestas: Ejemplo

Note que si extendemos Σ con dos nuevos átomos $Garrero(pepe)$ y $Bullanguero(pepe)$, entonces existe otra derivación de la cual obtenemos $pepe$ como respuesta.

Ejercicio: ¿Puede ser que obtengamos una cláusula con variables libres como respuesta a nuestra consulta? ¿Cómo debe interpretarse aquello?

Búsquedas más generales

Como dijimos anteriormente queremos mantener nuestras búsquedas (sustituciones, resoluciones, etc) lo más general posible.

Ejemplo: Sean átomos $P(x, z)$ y $P(b, y)$. Estos átomos se hacen iguales mediante las sutituciones:

- $x = b$ e $y = z$;
- $x = b$ e $y = z = b$.

¿Cuál de las dos sustituciones es más general y por qué?

Formalizaremos esta idea luego.

Unificación: Composición de sustituciones

Sean $\theta = \{x_1/t_1, \ldots, x_n/t_n\}$ y $\gamma = \{y_1/s_1, \ldots, y_m/s_m\}$ sustituciones.

La composición de θ y γ, denotada por $\theta\gamma$, es la sustitución que se obtiene del conjunto

$$\{x_1/t_1\gamma, \ldots, x_n/t_n\gamma, y_1/s_1, \ldots, y_m/s_m\}$$

al borrar cada ligazón $x_i/t_i\gamma$ tal que $t_i\gamma = x_i$, y cada ligazón y_j/s_j tal que $y_j \in \{x_1, \ldots, x_n\}$.

Unificación: Composición de sustituciones

Sean $\theta = \{x_1/t_1, \ldots, x_n/t_n\}$ y $\gamma = \{y_1/s_1, \ldots, y_m/s_m\}$ sustituciones.

La composición de θ y γ, denotada por $\theta\gamma$, es la sustitución que se obtiene del conjunto

$$\{x_1/t_1\gamma, \ldots, x_n/t_n\gamma, y_1/s_1, \ldots, y_m/s_m\}$$

al borrar cada ligazón $x_i/t_i\gamma$ tal que $t_i\gamma = x_i$, y cada ligazón y_j/s_j tal que $y_j \in \{x_1, \ldots, x_n\}$.

Ejemplo: Sea $\theta = \{x/f(y), y/z, w/u\}$ y $\gamma = \{x/a, y/b, z/y\}$.

Unificación: Composición de sustituciones

Sean $\theta = \{x_1/t_1, \ldots, x_n/t_n\}$ y $\gamma = \{y_1/s_1, \ldots, y_m/s_m\}$ sustituciones.

La composición de θ y γ, denotada por $\theta\gamma$, es la sustitución que se obtiene del conjunto

$$\{x_1/t_1\gamma, \ldots, x_n/t_n\gamma, y_1/s_1, \ldots, y_m/s_m\}$$

al borrar cada ligazón $x_i/t_i\gamma$ tal que $t_i\gamma = x_i$, y cada ligazón y_j/s_j tal que $y_j \in \{x_1, \ldots, x_n\}$.

Ejemplo: Sea $\theta = \{x/f(y), y/z, w/u\}$ y $\gamma = \{x/a, y/b, z/y\}$.

Entonces $\theta\gamma = \{x/f(b), w/u, z/y\}$.

Unificación: Propiedades de la composición

De ahora en adelante llamamos la susutitución identidad a la sustitución vacía $\{\}$. La denotamos por ϵ.

Ejercicio: Demuestre las siguientes propiedades de la composición de sustituciones:

- Para toda sustitución θ, $\theta\epsilon = \epsilon\theta = \epsilon$.
- Para toda expresión E y sustituciones θ y γ, $(E\theta)\gamma = E\theta\gamma$.
- Para toda sustituciones θ, γ, y ν, $(\theta\gamma)\nu = \theta(\gamma\nu)$.

Ejemplo: Considere $E = P(x, y, g(z))$, y θ y γ según nuestro último ejemplo.

Entonces $E\theta = P(f(y), z, g(z))$, y $(E\theta)\gamma = P(f(b), y, g(y))$.

Por el otro lado, $E\theta\gamma = P(f(b), y, g(y))$.

Unificación de expresiones

Sea $\mathcal{S}$ un conjunto *finito* de expresiones.

Una sustitución θ es un unificador de $\mathcal{S}$ si para todo $E, E' \in \mathcal{S}$ se tiene que $E\theta = E'\theta$.

Decimos que $\mathcal{S}$ es unificable si existe un unificador de $\mathcal{S}$.

Unificación de expresiones

Sea $\mathcal{S}$ un conjunto *finito* de expresiones.

Una sustitución θ es un unificador de $\mathcal{S}$ si para todo $E, E' \in \mathcal{S}$ se tiene que $E\theta = E'\theta$.

Decimos que $\mathcal{S}$ es unificable si existe un unificador de $\mathcal{S}$.

Ejemplo: Sea $\mathcal{S} = \{P(f(x), z), P(y, a)\}$.

Unificación de expresiones

Sea $\mathcal{S}$ un conjunto *finito* de expresiones.

Una sustitución θ es un unificador de $\mathcal{S}$ si para todo $E, E' \in \mathcal{S}$ se tiene que $E\theta = E'\theta$.

Decimos que $\mathcal{S}$ es unificable si existe un unificador de $\mathcal{S}$.

Ejemplo: Sea $\mathcal{S} = \{P(f(x), z), P(y, a)\}$.

$\mathcal{S}$ es unificable: $\theta = \{y/f(x), z/a\}$, $\theta' = \{y/f(a), x/a, z/a\}$, ...

Unificación de expresiones

Sea $\mathcal{S}$ un conjunto *finito* de expresiones.

Una sustitución θ es un unificador de $\mathcal{S}$ si para todo $E, E' \in \mathcal{S}$ se tiene que $E\theta = E'\theta$.

Decimos que $\mathcal{S}$ es unificable si existe un unificador de $\mathcal{S}$.

Ejemplo: Sea $\mathcal{S} = \{P(f(x), z), P(y, a)\}$.

$\mathcal{S}$ es unificable: $\theta = \{y/f(x), z/a\}$, $\theta' = \{y/f(a), x/a, z/a\}$, ...

Ejemplo: Sea $\mathcal{S} = \{P(f(x), a), P(y, f(w))\}$.

Unificación de expresiones

Sea $\mathcal{S}$ un conjunto *finito* de expresiones.

Una sustitución θ es un unificador de $\mathcal{S}$ si para todo $E, E' \in \mathcal{S}$ se tiene que $E\theta = E'\theta$.

Decimos que $\mathcal{S}$ es unificable si existe un unificador de $\mathcal{S}$.

Ejemplo: Sea $\mathcal{S} = \{P(f(x), z), P(y, a)\}$.

$\mathcal{S}$ es unificable: $\theta = \{y/f(x), z/a\}$, $\theta' = \{y/f(a), x/a, z/a\}$, ...

Ejemplo: Sea $\mathcal{S} = \{P(f(x), a), P(y, f(w))\}$.

$\mathcal{S}$ no es unificable: No hay forma de que $a = f(w)$.

Unificación de expresiones: Unificador más general

Presentaremos un algoritmo que toma un conjunto $\mathcal{S}$ y entrega un unificador θ si el conjunto es unificable.

El algoritmo no sólo entrega un unificador θ sino que θ es además el unificador más general (umg): Para cualquier otro unificador γ se tiene que existe ν tal que $\gamma = \theta\nu$.

Ejemplo: Sean átomos $P(x, z)$ y $P(b, y)$. Demuestre que la sustitución $\{x/b, y/b, z/b\}$ no es un umg.

Unificador más general

Pregunta: ¿Por qué nos interesa usar el umg?

Porque de otra forma podemos perder completitud en la resolución.

Después nos aseguraremos que para cierto tipo de cláusulas al utilizar el umg no se produce este tipo de problemas.

Unificación de expresiones: Intuición del algoritmo

El algoritmo funciona intuitivamente de la siguiente manera:

- El algoritmo se inicializa con punteros en el símbolo más a la izquierda de cada expresión $\mathcal{S}$.
- Los punteros se mueven sincronizadamente hacia la derecha de cada expresión hasta que apuntan a subtérminos diferentes.
- Se intenta unificar estas dos subexpresiones mediante alguna sustitución:
  - Si las subexpresiones son unificables entonces se sigue el proceso a partir de esta unificación.
  - Si no lo son, entonces $\mathcal{S}$ no es unificable y el algortimo falla.
- Si el algoritmo alcanza el fin de cada expresión en $\mathcal{S}$ entonces la composición de cada sustitución realizada por uno de los pasos del algortimo debe ser un umg.

Unificación de expresiones: Conjunto desacuerdo

Antes de definir el algoritmo es necesario definir el concepto de conjunto desacuerdo para un conjunto $\mathcal{S}$ de expresiones.

Primero, encuentre la posición del símbolo más a la izquierda para el cual no todas las expresiones en S tienen el mismo símbolo.

Extraiga de cada expresión en S la subexpresión que empieza en esa posición.

El conjunto desacuerdo es el conjunto de todas las subexpresiones extraidas en el paso anterior.

Unificación de expresiones: Conjunto desacuerdo

Antes de definir el algoritmo es necesario definir el concepto de conjunto desacuerdo para un conjunto $\mathcal{S}$ de expresiones.

Primero, encuentre la posición del símbolo más a la izquierda para el cual no todas las expresiones en S tienen el mismo símbolo.

Extraiga de cada expresión en S la subexpresión que empieza en esa posición.

El conjunto desacuerdo es el conjunto de todas las subexpresiones extraidas en el paso anterior.

Ejemplo: $\mathcal{S} = \{P(f(x), h(y), a), P(f(x), z, a), P(f(x), h(y), b)\}$.

Unificación de expresiones: Conjunto desacuerdo

Antes de definir el algoritmo es necesario definir el concepto de conjunto desacuerdo para un conjunto $\mathcal{S}$ de expresiones.

Primero, encuentre la posición del símbolo más a la izquierda para el cual no todas las expresiones en S tienen el mismo símbolo.

Extraiga de cada expresión en S la subexpresión que empieza en esa posición.

El conjunto desacuerdo es el conjunto de todas las subexpresiones extraidas en el paso anterior.

Ejemplo: $\mathcal{S} = \{P(f(x), h(y), a), P(f(x), z, a), P(f(x), h(y), b)\}$.

Entonces el conjunto desacuerdo de $\mathcal{S}$ es $\{h(y), z\}$.

Unificación de expresiones: Algoritmo

Algoritmo de unificación de expresiones:

Dado conjunto $\mathcal{S}$ de expresiones:

1. Inicialización: $k := 0$ y $\theta_0 = \epsilon$;
2. si θ_k unifica $\mathcal{S}$, pare; θ_k es umg de $\mathcal{S}$. Si no, encuentre el conjunto de desacuerdos D_k de $\mathcal{S}$;
3. si existen variable x y término t en D_k tal que t no menciona a x, entonces haga $\theta_{k+1} = \theta_k\{x/t\}$; incremente k y vaya a paso 2. Si no, $\mathcal{S}$ no es unificable.

Ejemplo: Sea $\mathcal{S} = \{P(f(a), g(x)), P(y, y)\}$.

Ejemplos de aplicación de algoritmo

Ejemplo: Sea $\mathcal{S} = \{P(f(a), g(x)), P(y, y)\}$.

Entonces,

1. $\theta_0 = \epsilon$;
2. $D_0 = \{f(a), y\}$, $\theta_1 = \{y/f(a)\}$,
   $\mathcal{S}\theta_1 = \{P(f(a), g(x)), P(f(a), f(a))\}$;
3. $D_1 = \{g(x), f(a)\}$, el algoritmo falla y $\mathcal{S}$ no es unificable.

Ejemplos de aplicación de algoritmo

Ejemplo: Sea $\mathcal{S} = \{P(f(a), g(x)), P(y, y)\}$.

Entonces,

1. $\theta_0 = \epsilon$;
2. $D_0 = \{f(a), y\}$, $\theta_1 = \{y/f(a)\}$,
   $\mathcal{S}\theta_1 = \{P(f(a), g(x)), P(f(a), f(a))\}$;
3. $D_1 = \{g(x), f(a)\}$, el algoritmo falla y $\mathcal{S}$ no es unificable.

Ejemplos de aplicación de algoritmo

Ejemplo: Sea $\mathcal{S} = \{P(f(a), g(x)), P(y, y)\}$.

Entonces,

1. $\theta_0 = \epsilon$;
2. $D_0 = \{f(a), y\}$, $\theta_1 = \{y/f(a)\}$,
   $\mathcal{S}\theta_1 = \{P(f(a), g(x)), P(f(a), f(a))\}$;
3. $D_1 = \{g(x), f(a)\}$, el algoritmo falla y $\mathcal{S}$ no es unificable.

Ejemplos de aplicación de algoritmo

Ejemplo: Sea $\mathcal{S} = \{P(a, x, h(g(z))), P(z, h(y), h(y))\}$.

Ejemplos de aplicación de algoritmo

Ejemplo: Sea $\mathcal{S} = \{P(a, x, h(g(z))), P(z, h(y), h(y))\}$.

Entonces,

1. $\theta_0 = \epsilon$;
2. $D_0 = \{a, z\}$, $\theta_1 = \{z/a\}$,
   $\mathcal{S}\theta_1 = \{P(a, x, h(g(a))), P(a, h(y), h(y))\}$;
3. $D_1 = \{x, h(y)\}$, $\theta_2 = \{z/a, x/h(y)\}$,
   $\mathcal{S}\theta_2 = \{P(a, h(y), h(g(a))), P(a, h(y), h(y))\}$;
4. $D_2 = \{y, g(a)\}$, $\theta_3 = \{z/a, x/h(y), y/g(a)\}$,
   $\mathcal{S}\theta_3 = \{P(a, h(g(a)), h(g(a)))$, $\mathcal{S}$ es unificable con umg θ_3.

Ejemplos de aplicación de algoritmo

Ejemplo: Sea $\mathcal{S} = \{P(a, x, h(g(z))), P(z, h(y), h(y))\}$.

Entonces,

1. $\theta_0 = \epsilon$;
2. $D_0 = \{a, z\}$, $\theta_1 = \{z/a\}$,
   $\mathcal{S}\theta_1 = \{P(a, x, h(g(a))), P(a, h(y), h(y))\}$;
3. $D_1 = \{x, h(y)\}$, $\theta_2 = \{z/a, x/h(y)\}$,
   $\mathcal{S}\theta_2 = \{P(a, h(y), h(g(a))), P(a, h(y), h(y))\}$;
4. $D_2 = \{y, g(a)\}$, $\theta_3 = \{z/a, x/h(y), y/g(a)\}$,
   $\mathcal{S}\theta_3 = \{P(a, h(g(a)), h(g(a)))$, $\mathcal{S}$ es unificable con umg θ_3.

Ejemplos de aplicación de algoritmo

Ejemplo: Sea $\mathcal{S} = \{P(a, x, h(g(z))), P(z, h(y), h(y))\}$.

Entonces,

1. $\theta_0 = \epsilon$;
2. $D_0 = \{a, z\}$, $\theta_1 = \{z/a\}$,
   $\mathcal{S}\theta_1 = \{P(a, x, h(g(a))), P(a, h(y), h(y))\}$;
3. $D_1 = \{x, h(y)\}$, $\theta_2 = \{z/a, x/h(y)\}$,
   $\mathcal{S}\theta_2 = \{P(a, h(y), h(g(a))), P(a, h(y), h(y))\}$;
4. $D_2 = \{y, g(a)\}$, $\theta_3 = \{z/a, x/h(y), y/g(a)\}$,
   $\mathcal{S}\theta_3 = \{P(a, h(g(a)), h(g(a)))$, $\mathcal{S}$ es unificable con umg θ_3.

Ejemplos de aplicación de algoritmo

Ejemplo: Sea $\mathcal{S} = \{P(a, x, h(g(z))), P(z, h(y), h(y))\}$.

Entonces,

1. $\theta_0 = \epsilon$;
2. $D_0 = \{a, z\}$, $\theta_1 = \{z/a\}$,
   $\mathcal{S}\theta_1 = \{P(a, x, h(g(a))), P(a, h(y), h(y))\}$;
3. $D_1 = \{x, h(y)\}$, $\theta_2 = \{z/a, x/h(y)\}$,
   $\mathcal{S}\theta_2 = \{P(a, h(y), h(g(a))), P(a, h(y), h(y))\}$;
4. $D_2 = \{y, g(a)\}$, $\theta_3 = \{z/a, x/h(y), y/g(a)\}$,
   $\mathcal{S}\theta_3 = \{P(a, h(g(a)), h(g(a)))$, $\mathcal{S}$ es unificable con umg θ_3.

Ejemplos de aplicación de algoritmo

Otro ejemplo que muestra por qué en el paso 3 t no puede mencionar a x.

Ejemplo: Sea $\mathcal{S} = \{P(x, x), P(y, f(y))\}$.

Otro ejemplo que muestra por qué en el paso 3 t no puede mencionar a x.

Ejemplo: Sea $\mathcal{S} = \{P(x,x), P(y, f(y))\}$.

Entonces,

1. $\theta_0 = \epsilon$;
2. $D_0 = \{x, y\}$, $\theta_1 = \{x/y\}$, $\mathcal{S}\theta_1 = \{P(x,x), P(x, f(x))\}$;
3. $D_1 = \{x, f(x)\}$, el algoritmo falla y $\mathcal{S}$ no es unificable.

Ejemplos de aplicación de algoritmo

Otro ejemplo que muestra por qué en el paso 3 t no puede mencionar a x.

Ejemplo: Sea $\mathcal{S} = \{P(x,x), P(y, f(y))\}$.

Entonces,

1. $\theta_0 = \epsilon$;
2. $D_0 = \{x, y\}$, $\theta_1 = \{x/y\}$, $\mathcal{S}\theta_1 = \{P(x,x), P(x, f(x))\}$;
3. $D_1 = \{x, f(x)\}$, el algoritmo falla y $\mathcal{S}$ no es unificable.

Otro ejemplo que muestra por qué en el paso 3 t no puede mencionar a x.

Ejemplo: Sea $\mathcal{S} = \{P(x,x), P(y, f(y))\}$.

Entonces,

1. $\theta_0 = \epsilon$;
2. $D_0 = \{x, y\}$, $\theta_1 = \{x/y\}$, $\mathcal{S}\theta_1 = \{P(x,x), P(x, f(x))\}$;
3. $D_1 = \{x, f(x)\}$, el algoritmo falla y $\mathcal{S}$ no es unificable.

Correctitud y completidad del algortimo

El siguiente teorema demuestra que el algoritmo es correcto y completo:

Teorema (Robinson)

Sea $\mathcal{S}$ un conjunto de expresiones. Si $\mathcal{S}$ es unificable entonces el algoritmo anterior termina y entrega un umg. Si no es unificable entonces el algoritmo así lo determina en paso 3.

Considere el conjunto

$$\mathcal{S} \; = \; \{P(x_1, \ldots, x_n),\; P(f(x_0, x_0), f(x_1, x_1), \ldots, f(x_{n-1}, x_{n-1}))\}.$$

Entonces:

- $\theta_1 = \{x_1/f(x_0, x_0)\}$ y
  $$\begin{aligned}\mathcal{S}\theta_1 = \{&P(f(x_0, x_0), \ldots, x_n),\\ &P(f(x_0, x_0),\; f(f(x_0, x_0), f(x_0, x_0)), \ldots, f(x_{n-1}, x_{n-1}))\}.\end{aligned}$$
- $\theta_2 = \{x_1/f(x_0, x_0),\; x_2/f(f(x_0, x_0), f(x_0, x_0))\}$ y
  $$\begin{aligned}\mathcal{S}\theta_2 = \{&P(f(x_0, x_0),\; f(f(x_0, x_0), f(x_0, x_0)), \ldots, x_n),\\ &P(f(x_0, x_0),\; f(f(x_0, x_0), f(x_0, x_0)),\\ &f(f(f(x_0, x_0), f(x_0, x_0)), f(f(x_0, x_0), f(x_0, x_0))), \ldots, f(x_{n-1}, x_{n-1}))\}.\end{aligned}$$
- .....

Complejidad del algortimo

Considere el conjunto

$$\mathcal{S} \;=\; \{P(x_1,\ldots,x_n),\; P(f(x_0,x_0), f(x_1,x_1),\ldots,f(x_{n-1},x_{n-1}))\}.$$

Entonces:

- $\theta_1 = \{x_1/f(x_0,x_0)\}$ y
$$\mathcal{S}\theta_1 = \{P(f(x_0,x_0),\ldots,x_n),\\ P(f(x_0,x_0),\; f(f(x_0,x_0),f(x_0,x_0)),\ldots,f(x_{n-1},x_{n-1}))\}.$$
- $\theta_2 = \{x_1/f(x_0,x_0),\; x_2/f(f(x_0,x_0),f(x_0,x_0))\}$ y
$$\mathcal{S}\theta_2 = \{P(f(x_0,x_0),\; f(f(x_0,x_0),f(x_0,x_0)),\ldots,x_n),\\ P(f(x_0,x_0),\; f(f(x_0,x_0),f(x_0,x_0)),\\ f(f(f(x_0,x_0),f(x_0,x_0)),f(f(x_0,x_0),f(x_0,x_0))),\ldots,f(x_{n-1},x_{n-1}))\}.$$
- .....

En particular, el k-esimo atributo del segundo átomo de $\mathcal{S}$ tendrá ocurrencias de f.

Considere el conjunto

$$\mathcal{S} \;=\; \{P(x_1, \ldots, x_n),\; P(f(x_0, x_0), f(x_1, x_1), \ldots, f(x_{n-1}, x_{n-1}))\}.$$

Entonces:

- $\theta_1 = \{x_1/f(x_0, x_0)\}$ y
$$\begin{aligned}\mathcal{S}\theta_1 = \{&P(f(x_0, x_0), \ldots, x_n),\\ &P(f(x_0, x_0),\; f(f(x_0, x_0), f(x_0, x_0)), \ldots, f(x_{n-1}, x_{n-1}))\}.\end{aligned}$$
- $\theta_2 = \{x_1/f(x_0, x_0),\; x_2/f(f(x_0, x_0), f(x_0, x_0))\}$ y
$$\begin{aligned}\mathcal{S}\theta_2 = \{&P(f(x_0, x_0),\; f(f(x_0, x_0), f(x_0, x_0)), \ldots, x_n),\\ &P(f(x_0, x_0),\; f(f(x_0, x_0), f(x_0, x_0)),\\ &f(f(f(x_0, x_0), f(x_0, x_0)), f(f(x_0, x_0), f(x_0, x_0))), \ldots, f(x_{n-1}, x_{n-1}))\}.\end{aligned}$$
- .....

En particular, el k-esimo atributo del segundo átomo de $\mathcal{S}$ tendrá $2^k - 1$ ocurrencias de f.

Complejidad del algortimo

Por tanto, el paso 3 del algoritmo toma tiempo exponencial.

Existen algoritmos de unificación más complejos pero de complejidad lineal.

PROLOG implementa el algoritmo que presentamos pero simplemente elimina el chequear si t menciona a x.

Como hemos visto, esto puede generar problemas en algunos casos.

¿Y si tenemos igualdad?

Una de las restricciones del teorema de Herbrand es que las cláusulas no contengan literales de la forma $t_1 = t_2$ o $\neg(t_1 = t_2)$.

El teorema de Herbrand relaciona el contenido semántico con el sintáctico.

Sin embargo, la igualdad es un tipo especial de relación y requiere tratamiento particular.

Para ello debemos explicarle al procedimiento de resolución como *entender* la igualdad.

Ejercicio: ¿Cómo podríamos *explicarle* al procedimiento que la cláusula

$$\{a = b,\ b = c,\ \neg(a = c)\}$$

es insatisfacible?

Axiomatización de la igualdad

Para que todas las propiedades de la igualdad sean tomadas en cuenta debemos incluir en nuestra base de conocimiento el siguiente conjunto $AX_=$ de cláusulas:

- (reflexividad) $x = x$.
- (simetría) $x = y \to y = x$
- (transitividad) $x = y \wedge y = z \to x = z$
- (sustitución de funciones) Para toda función n-aria f,

$$\forall x_1, \ldots, x_n (\bigwedge_{1 \leq i \leq n} x_i = y_i \ \to \ f(x_1, \ldots, x_n) = f(y_1, \ldots, y_n))$$

- (sustitución de relaciones) Para toda relación m-aria R,

$$\forall x_1, \ldots, x_n (\bigwedge_{1 \leq i \leq n} x_i = y_i \wedge R(x_1, \ldots, x_n) \ \to \ R(y_1, \ldots, y_n))$$

Resolución e igualdad

Es posible demostrar que si Σ es cualquier tipo de cláusulas (pueden mencionar igualdad), entonces:

Σ es insatisfacible si y sólo si la cláusula vacía puede ser obtenida mediante resolución desde $\Sigma \cup AX_=$.

Ejercicio: Demuestre usando resolución que la consulta $P(a, g(b))$ es consecuencia lógica de la base de conocimiento:

$$P(f(x), g(x)) \qquad f(b) = a$$

Paramodulación

El problema con las resoluciones que incluyen los axiomas de igualdad es que suelen ser muy largas y con un alto número de resolventes.

Una técnica que intenta hacer esto más eficiente es la llamada regla de paramodulación:

Si C_1 es una cláusula que menciona al termino t', y existe sustitución θ tal que $t\theta = t'\theta$, entonces de C_1 y $C_2 \cup \{s = t\}$ se puede inferir toda cláusula que se obtiene desde $(C_1 \cup C_2)\theta$ al reemplazar un subconjunto de las ocurrencias de t' por s.

Ejercicio: El ejercicio anterior pero ahora usando paramodulación.

Fórmulas no clausales

Ejercicio: Demuestre que para todo conjunto de oraciones Γ existe un conjunto Σ de cláusulas tal que

$$\Gamma \text{ es insatisfacible} \quad \Leftrightarrow \quad \Sigma \text{ es insatisfacible.}$$

Esto significa que cualquier conjunto de oraciones puede ser transformado en un conjunto *equiconsistente* de cláusulas, y luego se puede utilizar resolución para verificar su satisfacibilidad.

Observaciones finales

La resolución nos asegura que llegaremos a la cláusula vacía si Σ es insatisfacible.

Sin embargo, no nos asegura que el procedimiento va a terminar cuando Σ es satisfacible.

Incluso cuando Σ es satisfacible podemos equivocarnos y seguir demostraciones infinitas.

Esto tiene sentido: El problema de satisfacibilidad para la lógica de primer orden es indecidible.

Observaciones finales: No terminación

Ejemplo: No terminación.