

Paradigmas de Programación

Resolución SLD Prolog

1er cuatrimestre de 2024

Departamento de Computación
Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires

Resolución general

Recordemos

Para determinar si una fórmula de primer orden σ es válida:

1. Pasar su negación $\neg\sigma$ a forma clausal.

Se obtiene un conjunto \mathcal{C} de cláusulas tal que $\neg\sigma$ es satisfactible sii \mathcal{C} es satisfactible.

2. Aplicar repetidamente la regla de resolución:

$$\frac{\{\sigma_1, \dots, \sigma_p, \ell_1, \dots, \ell_n\} \quad \{\neg\tau_1, \dots, \neg\tau_q, \ell'_1, \dots, \ell'_m\} \quad (p, q > 0)}{\mathbf{S} = \text{mgu}(\{\sigma_1 \stackrel{?}{=} \sigma_2 \stackrel{?}{=} \dots \stackrel{?}{=} \sigma_p \stackrel{?}{=} \tau_1 \stackrel{?}{=} \tau_2 \stackrel{?}{=} \dots \stackrel{?}{=} \tau_q\})} \mathbf{S}(\{\ell_1, \dots, \ell_n, \ell'_1, \dots, \ell'_m\})$$

3. Si se alcanza la cláusula vacía, $\neg\sigma$ es insatisfactible y σ válida.
4. El método puede no terminar.

1

3

Resolución general

Aplicar el método de resolución general puede ser muy costoso.
Cada paso requiere usar criterios de búsqueda y selección:

Búsqueda. Elegir dos cláusulas.

Selección. Elegir un subconjunto de literales de cada cláusula.

La cantidad de opciones es exponencial en el tamaño del problema.

Además, cada paso agrega una nueva cláusula.

Además, en cada paso se deben resolver ecuaciones de unificación.

Además, el método requiere usar *breadth-first search* (BFS).

4

Resolución SLD

Veremos una variante de la resolución general, la **resolución SLD**.
Es un *tradeoff*: menor generalidad a cambio de mayor eficiencia.

Menor generalidad

No se puede aplicar sobre fórmulas de primer orden arbitrarias.
Sólo se puede aplicar sobre **cláusulas de Horn**.

Mayor eficiencia

Se reducen las opciones de búsqueda/selección.

5

Cláusulas de Horn

Recordemos que una cláusula es un conjunto de literales:

$$\{\ell_1, \dots, \ell_n\}$$

donde cada literal es una fórmula atómica posiblemente negada:

$$\ell ::= \underbrace{P(t_1, \dots, t_n)}_{\text{literal positivo}} \mid \underbrace{\neg P(t_1, \dots, t_n)}_{\text{literal negativo}}$$

Definición (Cláusulas de Horn)

Las cláusulas son de los siguientes tipos, dependiendo del número de literales positivos/negativos que contienen:

	#positivos	#negativos
cláusula objetivo	0	*
cláusula de definición	1	*
cláusula de Horn	≤ 1	*

Regla de resolución SLD

La regla de resolución SLD involucra siempre a una cláusula de **definición** y una cláusula **objetivo**:

$$\frac{\begin{array}{c} \{P(t_1, \dots, t_k), \neg\sigma_1, \dots, \neg\sigma_n\} \quad \{\neg P(s_1, \dots, s_k), \neg\tau_1, \dots, \neg\tau_m\} \\ S = \text{mgu}(\{P(t_1, \dots, t_k) \stackrel{?}{=} P(s_1, \dots, s_k)\}) \end{array}}{S(\{\neg\sigma_1, \dots, \neg\sigma_n, \neg\tau_1, \dots, \neg\tau_m\})}$$

Caso particular de la regla de resolución general.

La selección es binaria (un literal de cada cláusula).

La resolvente es una nueva cláusula **objetivo**.

Cláusulas de Horn

Ejemplo — cláusulas de Horn

$$\underbrace{\begin{array}{cc} \{P(X)\} & \{P(X), \neg Q(Y, X), \neg R(Y)\} \\ \text{cláusulas de } \textcolor{blue}{\text{definición}} & \end{array}}_{\text{cláusulas de } \textcolor{red}{\text{Horn}}} \quad \underbrace{\begin{array}{cc} \{\} & \{\neg P(X), \neg Q(Y, X)\} \\ \text{cláusulas } \textcolor{red}{\text{objetivo}} & \end{array}}$$

Observación

Hay fórmulas que no se pueden escribir como cláusulas de Horn.
Por ejemplo:

$$P \vee Q$$

6

Derivaciones SLD

Una derivación SLD comienza con $n \geq 0$ cláusulas de **definición** y una cláusula **objetivo**:

$$D_1 \quad \dots \quad D_n \quad G_1$$

En cada paso:

- ▶ Se elige una cláusula definición D_j con $1 \leq j \leq n$.
- ▶ Se aplica la regla de resolución SLD sobre D_j y G_i .
- ▶ La resolvente es una nueva cláusula objetivo G_{i+1} .

8

7

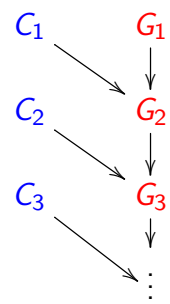
9

Derivaciones SLD

Es decir, dado un conjunto de cláusulas:

$D_1 \quad \dots \quad D_n \quad G_1$

una derivación SLD es de la forma:



Cada C_i debe ser alguna de las cláusulas originales $\{D_1, \dots, D_n\}$.
La cláusula G_{i+1} se obtiene aplicando resolución SLD sobre C_i y G_i .

Derivaciones SLD

Ejemplo

Dadas las hipótesis:

$\boxed{1} \quad \forall X. a(0, X, X)$
 $\boxed{2} \quad \forall X. \forall Y. \forall Z. (a(X, Y, Z) \Rightarrow a(s(X), Y, s(Z)))$

Queremos demostrar:

$\boxed{3} \quad \exists X. a(s(0), X, s(s(s(0))))$

Es decir, queremos probar que $(\boxed{1} \wedge \boxed{2}) \Rightarrow \boxed{3}$ es válida.
Basta ver que $\neg((\boxed{1} \wedge \boxed{2}) \Rightarrow \boxed{3})$ es insatisfactible.
Es decir, basta ver que $\boxed{1} \wedge \boxed{2} \wedge \neg \boxed{3}$ es insatisfactible.

Derivaciones SLD

Observaciones:

- La **búsqueda** se simplifica.
Se limita a elegir C_i como una de las n cláusulas D_1, \dots, D_n .
La cláusula objetivo G_i está fijada, no hay alternativas.
- La **selección** se simplifica.
Se limita a elegir uno de los literales **negativos** de G_i .
La cláusula de definición C_i tiene un único literal **positivo**.

Derivaciones SLD

Escribiendo $\boxed{1} \wedge \boxed{2} \wedge \neg \boxed{3}$ en forma clausal, tenemos:

$\boxed{1} \quad \{a(0, X, X)\}$
 $\boxed{2} \quad \{\neg a(X, Y, Z), a(s(X), Y, s(Z))\}$
 $\boxed{3} \quad \{\neg a(s(0), X, s(s(s(0))))\}$

$\boxed{1}$ y $\boxed{2}$ son cláusulas de **definición**.
 $\boxed{3}$ es la cláusula **objetivo**.

Busquemos una **refutación SLD** (derivación SLD que llega a $\{\}$):

$[3] = \{\neg a(s(0), X, s(s(s(0))))\}$

► $[3]$ y $[2] = \{\neg a(X_4, Y_4, Z_4), a(s(X_4), Y_4, s(Z_4))\}$

$S_4 = \{X_4 := 0, X := Y_4, Z_4 := s(s(0))\}.$

$[4] = \{\neg a(0, Y_4, s(s(0)))\}.$

► $[4]$ y $[1] = \{a(0, X_5, X_5)\}$

$S_5 = \{Y_4 := s(s(0)), X_5 := s(s(0))\}$

$[5] = \{\}$

Completitud del método de resolución SLD

El método de resolución es completo para cláusulas de Horn.

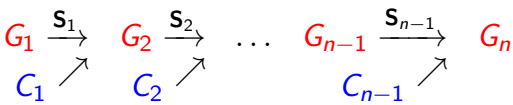
Más precisamente, si D_1, \dots, D_n son cláusulas de **definición** y G una cláusula **objetivo**:

Teorema

Si $\{D_1, \dots, D_n, G\}$ es insatisfactible, existe una refutación SLD.

Definición (Sustitución respuesta)

Dada una refutación SLD, con pasos:



la **sustitución respuesta** es la composición $S_{n-1} \circ \dots \circ S_1.$

Ejemplo — sustitución respuesta

En el ejemplo anterior, la sustitución respuesta es $S_5 \circ S_4.$

El valor de X en la cláusula objetivo original $[3]$ es $s(s(0)).$

Esto dice que $X = s(s(0))$ verifica nuestra consulta original:

$\exists X. a(s(0), X, s(s(s(0))))$

Semántica de Prolog

Un programa en Prolog es una **lista** de cláusulas de definición.

Una cláusula es también una lista. Una consulta en Prolog es una cláusula objetivo.

La notación cambia ligeramente.

Ejemplo

Cláusulas	Prolog
$\{a(0, X, X)\}$	$a(0, X, X) .$
$\{a(s(X), Y, s(Z)), \neg a(X, Y, Z)\}$	$a(s(X), Y, s(Z)) :- a(X, Y, Z) .$
$\{\neg a(s(0), X, s(s(s(0))))\}$	$?- a(s(0), X, s(s(s(0)))) .$

Semántica de Prolog

La ejecución se basa en la regla de resolución SLD.
Escrita en notación de Prolog:

$$\frac{?- p(t_1, \dots, t_k), \sigma_1, \dots, \sigma_n. \quad p(s_1, \dots, s_k) :- \tau_1, \dots, \tau_m.}{S = mgu(p(t_1, \dots, t_k) \stackrel{?}{=} p(s_1, \dots, s_k))} \\ S(?\tau_1, \dots, \tau_m, \sigma_1, \dots, \sigma_n.)$$

El orden de los literales en la cláusula objetivo es relevante.
Criterio de **selección**: elegir siempre el **primer literal** de la cláusula.

Semántica de Prolog

La exploración *depth-first* (DFS) es **incompleta**.
Puede provocar que Prolog nunca encuentre refutaciones posibles.

Ejemplo — incompletitud de DFS

```
esMaravilloso(X) :- esMaravilloso(suc(X)).
esMaravilloso(cero).

?- esMaravilloso(cero).
  |
  |_ ?- esMaravilloso(suc(cero)).
    |
    |_ ?- esMaravilloso(suc(suc(cero))).
      |
      |_ ?- esMaravilloso(suc(suc(suc(cero)))).
        |
        |_ ...
```

El orden de las reglas se torna relevante.

Tradeoff: mayor eficiencia a cambio de menor declaratividad.
La exploración *breadth-first* (BFS) es completa pero muy costosa.

Semántica de Prolog

Prolog busca sucesivamente todas las refutaciones haciendo DFS.
Criterio de **búsqueda**: las reglas se usan en orden de aparición.

Ejemplo — árbol de refutación

```
fruta(manzana).      dulce(manzana).
fruta(limón).         dulce(pera).
fruta(pera).
meGusta(helado).     meGusta(X) :- fruta(X), dulce(X).

?- meGusta(X).
  |
  |_ ✓ ..... {X := helado}
  |
  |_ ?- fruta(X), dulce(X).
    |
    |_ ?- dulce(manzana). ..... {X := manzana}
      |
      |_ ✓ ..... {X := manzana}
      |
      |_ ?- dulce(limón). ..... {X := limón}
        |
        |_ ✗
        |
        |_ ?- dulce(pera). ..... {X := pera}
          |
          |_ ✓ ..... {X := pera}
```

Semántica de Prolog

Al unificar, Prolog **no** usa la regla *occurs-check*.
Por ejemplo, X unifica con f(X). Esto es **incorrecto**.
Puede provocar que Prolog encuentre una “refutación” incorrecta.

Ejemplo — refutación incorrecta por omisión de *occurs check*

```
esElSucesor(X, suc(X)).

?- esElSucesor(Y, Y).
  |
  |_ ✓ ..... {Y := X, X := suc(X)}
```

Tradeoff: mayor eficiencia a cambio de incorrección lógica.
En muchos contextos la regla *occurs-check* es innecesaria.
La carga de probar corrección recae en lxs programadorxs.
Nota: En swi-prolog se puede cambiar usando
set_prolog_flag(occurs_check, true).

```
c([], Ys, Ys).
c([X | Xs], Ys, [X | Zs]) :- c(Xs, Ys, Zs).

?- c([1, 2], [3, 4], Zs).
  |_- ?- c([2], [3, 4], Zs1). .....{Zs := [1 | Zs1]}
  |   |_- ?- c([], [3, 4], Zs2). ..{Zs := [1, 2 | Zs2]}
  |       |_- ✓ .....{Zs := [1, 2, 3, 4]}

?- c([1, 2], Ys, [1, 2, 3, 4]).
  |_- ?- c([2], Ys, [2, 3, 4]).
  |   |_- ?- c([], Ys, [3, 4]).
  |       |_- ✓ .....{Ys := [3, 4]}
```

```
c([], Ys, Ys).
c([X | Xs], Ys, [X | Zs]) :- c(Xs, Ys, Zs).

?- c(Xs, [9], Zs).
  |_- ✓ .....{Xs := [], Zs := [9]}
  |   |_- ?- c(Xs1, [9], Zs1).
  |       |_- {Xs := [_1 | Xs1], Zs := [_1 | Xs1]}
  |           |_- ✓ .....{Xs := [_1], Zs := [_1, 9]}
  |               |_- ?- c(Xs2, [9], Zs2).
  |                   |_- {Xs := [_1, _2 | Xs2], Zs := [_1, _2 | Xs1]}
  |                       |_- ✓ .....{Xs := [_1, _2], Zs := [_1, _2, 9]}
  |                           |_- ...
```

```
c([], Ys, Ys).
c([X | Xs], Ys, [X | Zs]) :- c(Xs, Ys, Zs).

?- c(Xs, [3], [1, 2, 3]).
  |_- ?- c(Xs1, [3], [2, 3]). .....{Xs := [1 | Xs1]}
  |   |_- ?- c(Xs2, [3], [3]). .....{Xs := [1, 2 | Xs2]}
  |       |_- ✓ .....{Xs := [1, 2]}
  |           |_- ?- c(Xs3, [3], []).
  |               |_- {Xs := [1, 2, 3 | Xs3]}
  |                   |_- ✗
```

```
Consideremos el siguiente programa:

padre(zeus, atenea). % ...base de conocimiento...
ancestro(X, Y) :- padre(X, Y).
ancestro(X, Y) :- padre(X, Z), ancestro(Z, Y).

?- ancestro(zeus, atenea).
  |_- ?- padre(zeus, atenea).
  |   |_- ✓
  |   |_- ?- padre(zeus, Z), ancestro(Z, atenea).
  |       |_- ...
```

Nos gustaría tener una manera de podar el árbol de búsqueda.

Operador de corte (cut)

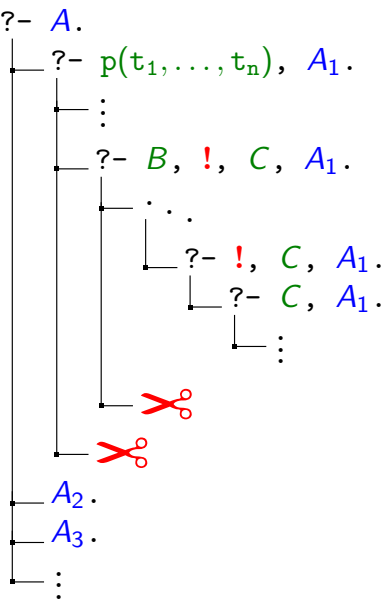
Reescribimos el programa agregando un *corte* ("!"):

```
padre(zeus, atenea). % ...base de conocimiento...
ancestro(X, Y) :- padre(X, Y), !.
ancestro(X, Y) :- padre(X, Z), ancestro(Z, Y).
```

El operador **!** no tiene una interpretación declarativa/lógica.
Es un operador extra-lógico.
Su comportamiento se explica desde el punto de vista operacional.
El operador de corte indica que, si se lo alcanza, no se deben explorar alternativas a la regla en la que aparece.

Operador de corte (cut)

Gráficamente:



Operador de corte (cut)

Semántica del operador de corte
El predicado **!** tiene éxito inmediatamente.

- Al momento de hacer *backtracking*:
- ▶ Se vuelve atrás hasta el punto en el que se eligió usar la regla que hizo aparecer el operador de corte.
 - ▶ Se descartan todas las elecciones alternativas.
 - ▶ Se continúa buscando hacia atrás.

Operador de corte (cut)

Ejemplo — cortes “benignos” (green cuts)
En algunos casos, **!** no altera la semántica del programa.
Puede servir para construir programas equivalentes más eficientes.

```
add(N, zero, N) :- !.
add(zero, N, N).
add(suc(N), M, suc(P)) :- add(N, M, P).
```

```
?- add(suc(suc(suc(suc(...))), zero, P).
```

Ejemplo — cortes “riesgosos” (red cuts)
En otros casos, la semántica puede verse alterada.
maximo(A, B, A) :- A >= B, !.
maximo(A, B, B).

?- maximo(2, 1, B).	?- maximo(2, 1, 1).
>> B = 2	>> true.

Negación por falla

Definición (operador de negación)

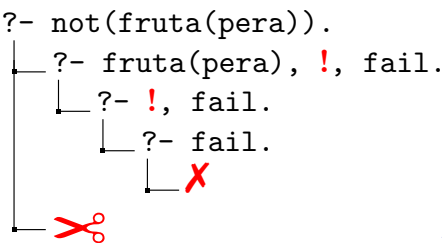
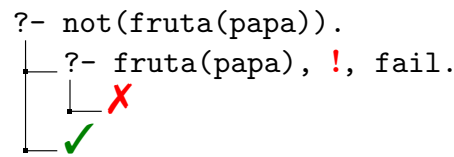
Si fail es un predicado que falla siempre, se puede definir la negación en Prolog así:

```
not(P) :- P, !, fail.  
not(P).
```

Observación. not(P) tiene éxito si y sólo si P falla.

Ejemplo — negación por falla

```
fruta(pera).
```



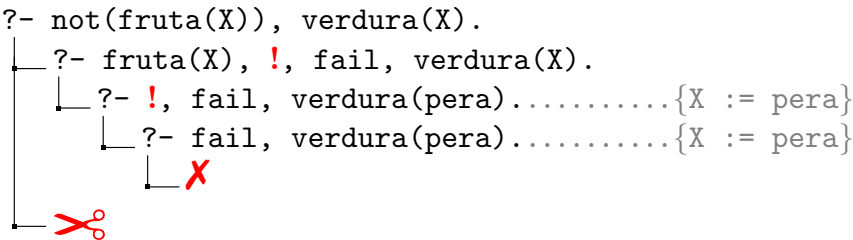
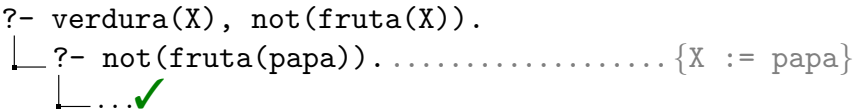
Ejemplo

Negación por falla

La negación por falla no coincide con la negación lógica.

Ejemplo — negación por falla ≠ negación lógica

```
fruta(pera).  
verdura(papa).
```



El orden de los literales en la consulta se vuelve relevante.
Esto atenta contra la declaratividad.

- 1. Definir append(L1, L2, L3).
- 2. Definir ocurreAlMenosUnaVez(X, L) usando append/3.
- 3. Definir ocurreAlMenosDosVeces(X, L) usando append/3.
- 4. Definir ocurreExactamenteUnaVez(X, L) combinando las anteriores.