TEMA IV

Modelos probabilísticos

Mario de J. Pérez Jiménez
Grupo de investigación en Computación Natural
Dpto. Ciencias de la Computación e Inteligencia Artificial
Universidad de Sevilla

Simulación y análisis computacional en Biología de Sistemas
Máster Universitario en Lógica, Computación e Inteligencia Artificial

Curso 2012-13

Modelo probabilístico basado en sistemas P (I)

Esqueleto de sistema P de grado $q \geq 1$:

$$\Pi = (\Gamma, \mu, \mathcal{R})$$

- Γ es un alfabeto (de trabajo);
- μ es una estructura de membranas: q membranas etiquetadas inyectivamente por $1, \ldots, q$ (1 etiqueta la membrana piel), con cargas eléctricas del conjunto $\{0, +, -\}$;
- $\mathcal{R}$ es un conjunto finito de reglas del tipo

  $$r: \; u[\, v \,]_h^{\alpha} \to u'[\, v' \,]_h^{\alpha'}$$

  en donde $u, v, u', v' \in \Gamma^*$, $h \in \{1, \ldots, q\}$, y $\alpha, \alpha' \in \{0, +, -\}$.

Es un conjunto de membranas polarizadas jerarquizado por una estructura μ. Inicialmente, todas las membranas están con carga neutra.

Modelo probabilístico basado en sistemas P (II)

Sistema P funcional de grado $q \geq 1$ y con $T \geq 1$ unidades de tiempo:

$$\Pi = (\Gamma, \mu, R, T, \{f_r : \ r \in R\}, \mathcal{M}_1, \ldots, \mathcal{M}_q)$$

- (Γ, μ, R) es el esqueleto de un sistema P extendido con membranas activas de grado q.
- T es un número natural;
- Para cada $r \in R$, f_r es una función computable tal que $dom(f_r) \subseteq \{1, \ldots, T\}$.
- $\mathcal{M}_1, \ldots, \mathcal{M}_q$ son multiconjuntos sobre Γ.

Es un esqueleto de sistema P con:

- Multiconjuntos iniciales en cada una de las membranas.
- Funciones computables asociadas a cada regla.
- Un tiempo máximo de ejecución T.

Si $r \in R$ y $t = 1, \ldots, T$, $f_r(t)$ es una constante asociada a r en el instante t. Notaremos $r : u[v]_i^{\alpha} \xrightarrow{f_r(t)} u'[v']_i^{\alpha'}$. Si $f_r(t) = 1$, omitiremos la constante: $r : u[v]_i^{\alpha} \longrightarrow u'[v']_i^{\alpha'}$.

Sistema P funcional:

Configuración del sistema en un instante t: tupla formada por los multiconjuntos presentes en las q *regiones* en t, junto con las polarizaciones.

Configuración inicial del sistema: $(\mathcal{M}_1, \dots, \mathcal{M}_q)$ y polarizaciones neutras.

Paso de transición de una configuración a otra: aplicando las reglas de R

- En la forma habitual
  - Una regla $u[v]_h^{\alpha} \rightarrow u'[v']_h^{\alpha'}$ es aplicable a una membrana etiquetada por h con carga α si u está contenido en la membrana padre de h y v está contenido en la membrana h con carga α.
- Pero de manera que la familia $\{f_r : \ r \in R\}$ materializa la dinámica o aplicabilidad de las reglas.

Modelo probabilístico basado en sistemas P (III)

Un sistema PDP (Population Dynamics P System) de grado (q, m) y con T unidades de tiempo $(q, m, T \geq 1)$ es una tupla:

$$(G, \Gamma, \Sigma, \mathcal{R}_\Pi, \mathcal{R}_E, \{f_{r,k} : \ r \in \mathcal{R}_\Pi, 1 \leq k \leq m\}, \{\mathcal{M}_{ik} : \ 1 \leq i \leq q, \ 1 \leq k \leq m\})$$

- $G = (V, S)$ es un grafo dirigido tal que $(e, e) \in S$ $(\forall e \in V)$. Los elementos de $V = \{e_1, \ldots e_m\}$ son los entornos.
- Γ es el alfabeto de trabajo y $\Sigma \subsetneq \Gamma$ (objetos que pueden estar presente en los entornos).
- $\Pi = (\Gamma, \mu, \mathcal{R}_\Pi)$ es el esqueleto de un sistema P extendido con membranas activas de grado q.
- Para cada k $(1 \leq k \leq m)$, $\Pi_k = (\Gamma, \mu, \mathcal{R}_\Pi, T, \{f_{r,k} : \ r \in \mathcal{R}_\Pi\}, \mathcal{M}_{1,k}, \ldots, \mathcal{M}_{q,k})$ es un sistema P funcional de grado q y con T unidades de tiempo tal que para cada $r \in \mathcal{R}_\Pi$ y k $(1 \leq k \leq m)$:
  - $f_{r,k}$ es una función computable cuyo rango está contenido en $[0, 1]$.
  - Para cada $u, v \in \Gamma^*$, $i \in \{1, \ldots, q\}$, $\alpha \in \{0, +, -\}$, si $r_1, \ldots, r_z$ son las reglas de $\mathcal{R}_\Pi$ cuya parte izquierda es $u \ [\ v \ ]_i^\alpha$, entonces $\sum_{j=1}^{z} f_{r_j,k}(t) = 1$, para cada t $(1 \leq t \leq T)$.
- $\mathcal{R}_E$ es un conjunto finito de reglas de comunicación de la forma $(x)_{e_j} \xrightarrow{p_{(x,j,j')}} (y)_{e_{j'}}$, donde $x, y \in \Sigma$, $(e_j, e_{j'}) \in S$, $1 \leq j, j' \leq m$, y $p_{(x,j,j')}$ son funciones computables cuyo dominio es $\{1, 2, \ldots, T\}$ y su rango está contenido en $[0, 1]$, de tal manera que si para e_j se tiene que $\{e_{j_1}, \ldots, e_{j_z}\}$ es el conjunto de nodos alcanzables desde e_j, entonces $\sum_{i=1}^{z} p_{(x,j,j_i)}(t) = 1$, para cada $x \in \Sigma$, $1 \leq t \leq T$
- Cada entorno e_k contiene un único sistema: Π_k.

Un sistema PDP de grado (q, m) y con T unidades de tiempo:

- Es un conjunto de m entornos conectados entre sí por los arcos de G.
- Cada entorno contiene un sistema Π_k, todos ellos con el mismo esqueleto.
- Las funciones $f_{r,k}(t)$ y $p_{x,j,j'}(t)$ proporcionan las probabilidades que tienen esas reglas de ser aplicadas en el instante t, en caso de ser aplicables.
- Cuando se aplica una regla del entorno del tipo

  $$(x)_{e_j} \xrightarrow{p_{(x,j,j')}} (y)_{e_{j'}}$$

  el objeto x pasa de e_j a $e_{j'}$ transformándose en y.
- La suma de todas las probabilidades de las reglas que tienen la misma *parte izquierda* es igual a 1.

El algoritmo DNDP (I)

Esquema algorítmico

Input: Un sistema PDP de grado (q, m), con T unidades de tiempo.

for $t \leftarrow 0$ to $T - 1$ **do**

$C_t \leftarrow$ configuración del sistema en el instante t

$C'_t \leftarrow C_t$

Inicialización

Primera fase de selección

(genera un multiconjunto *consistente* de reglas applicables)

Segunda fase de selección:

(genera un multiconjunto *maximal consistente* de reglas aplicables)

Ejecución de las reglas seleccionadas

$C_{t+1} \leftarrow C'_t$

end for

El algoritmo DNDP (II)

1. Inicialización

$R_\Pi \leftarrow$ conjunto ordenado de reglas de Π

for $j \leftarrow 1$ to m **do**

$R_{E,j} \leftarrow$ conjunto ordenado de reglas de R_E relativo al entorno e_j

$A_j \leftarrow$ conjunto ordenado de reglas de $R_{E,j}$ con probabilidad > 0 en el instante t

$M_j \leftarrow$ conjunto ordenado de pares $\langle label, charge \rangle$, para toda membrana de C_t contenida en el entorno e_j

$B_j \leftarrow \emptyset$

for each $\langle h, \alpha \rangle \in M_j$ (siguiendo el orden considerado) **do**

$B_j \leftarrow B_j \cup$ conjunto ordenado de reglas $u[v]_h^\alpha \leftarrow u'[v']_h^\beta$ de R_Π con probabilidad > 0 en t para e_j

end for

end for

El algoritmo DNDP (III)

2. Primera fase de selección (*consistencia*)

for $j \leftarrow 1$ to m **do**
  $R^1_{sel,j} \leftarrow \emptyset$
  $R^2_{sel,j} \leftarrow \emptyset$
  for $k \leftarrow 1$ to K **do**
    $D_j \leftarrow A_j \cup B_j$ con un *orden aleatorio*
    for each $r \in D_j$ (siguiendo el orden considerado) **do**
      if r es *consistente* con las reglas en $R^1_{sel,j}$ **then**
        $N' \leftarrow$ máx{número de veces que r es aplicable a C'_t}
        if $N' > 0$ **then**
          if $p_{r,j}(t) = 1$ **then**
            $n \leftarrow F_b(N', 0.5)$
          si no
            $N \leftarrow$ máx{numero de veces que r es aplicable a C_t}
            $n \leftarrow F_b(N, p_{r,j}(t))$
            if $n > N'$ **then**
              $n \leftarrow N'$
            end if
          end if
          if $n > 0$ **then**
            $C'_t \leftarrow C'_t - n \cdot l(r)$
            $R^1_{sel,j} \leftarrow R^1_{sel,j} \cup \{< r, n >\}$
          si no
            $R^2_{sel,j} \leftarrow R^2_{sel,j} \cup \{< r, n >\}$
          end if
        end if
      end if
    end for
  end for
end for

El algoritmo DNDP (IV)

3. Segunda fase de selección (*maximalidad*)

for $j \leftarrow 1$ to m **do**

$R_{sel,j} \leftarrow R^1_{sel,j} + R^2_{sel,j}$ (*orden decreciente en función de las probabilidades de las reglas*)

for each $< r, n > \in R_{sel,j}$ (siguiendo el orden seleccionado) **do**

if $n > 0 \vee$ (r es *consistente* con las reglas de $R^1_{sel,j}$) **then**

$N' \leftarrow \max\{\text{number of times that r is applicable to } C'_t\}$

if $N' > 0$ **then**

$R^1_{sel,j} \leftarrow R^1_{sel,j} \cup \{< r, N' >\}$

$C'_t \leftarrow C'_t - N' \cdot l(r)$

end if

end if

end for

end for

El algoritmo DNDP (V)

4. Ejecución de las reglas seleccionadas

for each $< r, n > \in R^1_{sel,j}$ **do**

$C'_t \leftarrow C'_t + n \cdot r(r)$

Actualizar las cargas eléctricas de C'_t de acuerdo con $r(r)$

end for

Observaciones

1. Dada una regla r, notaremos $l(r)$ la parte izquierda de dicha regla, y $r(r)$ la parte derecha.
2. Dos reglas r y r' con las mismas etiquetas y cargas eléctricas para $l(r)$ y $l(r')$ se dicen que son *consistentes* si $r(r)$ y $r(r')$tienen la *misma* carga eléctrica.

3 $p_{r,j}(t)$ indica la probabilidad asociada a la regla r del sistema P situado en el entorno j en el instante t.

4. Los multiconjuntos usados pueden contener pares ordenados $\langle r, 0\rangle$