

Grammaires de descriptions d'objets

Florent Hivert

Mél:Florent.Hivert@lri.fr
Adresse universelle:http://www.lri.fr/~hivert



Résumé des épisodes précédents

On a vu sur des exemples (mots binaires, permutations) plusieurs technique pour résoudre les questions classiques de la génération combinatoire :

- Énumération lexicographique;
- Décomposition récursive ;
- itération vs listage pour gagner en complexité.

Solution ad hoc ⇒ solution générique



Résumé des épisodes précédents

On a vu sur des exemples (mots binaires, permutations) plusieurs technique pour résoudre les questions classiques de la génération combinatoire :

- Énumération lexicographique;
- Décomposition récursive ;
- itération vs listage pour gagner en complexité.

Solution ad hoc ⇒ solution générique



Objectifs : algorithmes génériques

- Identifier les composants de base :
 - ⇒ Singleton, union, produit cartésien, ensemble et multiensemble. . .
- Comprendre comment composer les briques de base :
 - ⇒ grammaire de description, classe combinatoire
- ⇒ Algorithmes génériques!



Objectifs : algorithmes génériques

- Identifier les composants de base :
 - \Longrightarrow Singleton, union, produit cartésien, ensemble et multiensemble. . .
- Comprendre comment composer les briques de base :
 - ⇒ grammaire de description, classe combinatoire
- ⇒ Algorithmes génériques!

Union disjointe 4 de 35



Union disjointe

Definition

On écrit $C = A \sqcup B$ et on dit que C est l'union disjointe de A et B si $C = A \sqcup B$ et $A \cap B = \emptyset$.

Alors

- \blacksquare count(C) = count(A) + count(B)
- On peut prendre : list(C) = concat(list(A), list(B))

Union disjointe 4 de 35



Union disjointe

Definition

On écrit $C = A \sqcup B$ et on dit que C est l'union disjointe de A et B si $C = A \sqcup B$ et $A \cap B = \emptyset$.

Alors:

- \blacksquare count(C) = count(A) + count(B)
- On peut prendre : list(C) = concat(list(A), list(B))





Itération sur une union disjointe

On fixe l'ordre d'énumération tel que

$$list(A \sqcup B) := concat(list(A), list(B))$$

Itération en Python :

```
1 def iterunion(A, B):
2 for a in A:
3 yield a
4 for b in B:
5 yield b
```

Union disjointe

6 de 35

SUD

first, next sur une union disjointe

```
list(A \sqcup B) := concat(list(A), list(B))
1
          def first_union(A, B):
              return A.first()
3
          def next_union(A, B, x):
4
5
              if x in A:
6
                   try:
                       return A.next(x)
8
                   except StopIteration:
                       return B.first()
9
10
              else:
11
                   return B.next(x)
```

Union disjointe 7 de 35

rank sur une union disjointe

```
list(A \sqcup B) := concat(list(A), list(B))
 1
         def rank_union(A, B, x):
              if x in A:
3
                  return A.rank(x)
              else:
4
 5
                   return A.count() + B.rank(x)
6
         def unrank_union(A, B, i):
              if i < A.count():
                  return A.unrank(i)
9
10
              else:
                   return B.unrank(i - A.count())
11
```

ARIS Union disjointe 8 de 35

Le principe de l'idée récursive

Quand on a un'bonne idée, on l'appliqu'récursivement : on obtient le plus souvent une bien meilleure idée!

Unions disjointes récursives

Union disjointe 8 de 35

Le principe de l'idée récursive

Quand on a un'bonne idée, on l'appliqu'récursivement : on obtient le plus souvent une bien meilleure idée!

Unions disjointes récursives

Union disjointe 9 de 35

Les chaînes de *n*-bits ayant *k*-bits à 1

Une chaîne de bit non vide commence soit par un 0, soit par un 1 :

$$BitString(n, k) = 0 \cdot BitString(n-1, k) \sqcup 1 \cdot BitString(n-1, k-1)$$

Idem triangle de pascal:

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$$

■ BitString(n, k). count $() = \binom{n}{k}$

rank, unrank pour les chaînes de n-bits ayant k-bits à 1

```
1
     def rank BSnk(x):
         if not x:
                          # liste vide
 3
             return 0
         if x[0] == 0:
4
 5
             return rank_BSnk(x[1:])
6
         else:
 7
             return binom(len(x)-1, sum(x)-1) + rank_BSnk(x[1:])
8
9
     def unrank_BSnk(n, k, i):
10
         if n == 0:
11
             return []
12
         bn1k = binom(n-1, k)
13
         if i < bn1k:
14
             return [0]+unrank_BSnk(n-1, k, i)
15
         else:
16
             return [1]+unrank_BSnk(n-1, k-1, i-bn1k)
```

11 de 35



Le problème du calcul de la cardinalité

Problème

Le calcul récursif des coefficients binomiaux $\binom{n}{k}$ n'est pas efficace car on recalcule plusieurs fois la même chose.

Plus généralement, le calcul récursif des cardinalités sera très inefficace pour la même raison.

Union disjointe 12 de 35

Parenthèse : mémoization et programmation dynamique

Retenir

- **Mémoisation** : on mémorise tous les calculs pendant la récursion au momment où on les fait
- Programmation Dynamique : résoud les sous-problèmes, des plus petits aux plus grands en stockant les résultats intermédiaires.

En général, la programmation dynamique est plus efficace mais plus longue à mettre en oeuvre : il faut avoir planifié l'utilisation de la mémoire

Union disjointe 12 de 35

Parenthèse : mémoization et programmation dynamique

Retenir

- **Mémoisation** : on mémorise tous les calculs pendant la récursion au momment où on les fait
- **Programmation Dynamique**: résoud les sous-problèmes, des plus petits aux plus grands en stockant les résultats intermédiaires.

En général, la programmation dynamique est plus efficace mais plus longue à mettre en oeuvre : il faut avoir planifié l'utilisation de la mémoire.

Union disjointe 12 de 35

Parenthèse : mémoization et programmation dynamique

Retenir

- **Mémoisation** : on mémorise tous les calculs pendant la récursion au momment où on les fait
- **Programmation Dynamique**: résoud les sous-problèmes, des plus petits aux plus grands en stockant les résultats intermédiaires.

En général, la programmation dynamique est plus efficace mais plus longue à mettre en oeuvre : il faut avoir planifié l'utilisation de la mémoire.

Union disjointe 13 de 35

Autre exemple : les permutations

Les permutés d'un ensemble $X := \{x_1, x_2, \dots, x_n\}$:

 $\mathsf{Perm}\{1,2,3\} = 1 \cdot \mathsf{Perm}\{2,3\} \ \sqcup \ 2 \cdot \mathsf{Perm}\{1,3\} \ \sqcup \ 3 \cdot \mathsf{Perm}\{1,2\}$

Plus généralement :

Retenir

Énumération lexicographique des permutations :

$$\operatorname{Perm}(X) = \bigsqcup_{i=1}^{n} x_i \cdot \operatorname{Perm}(X/\{x_i\})$$

■ Perm(X). count() = |X|!

Généralisation : permuté d'un multiensemble

 $\mathsf{Perm}\{1,1,2,3\} = 1 \cdot \mathsf{Perm}\{1,2,3\} \sqcup 2 \cdot \mathsf{Perm}\{1,1,3\} \sqcup 3 \cdot \mathsf{Perm}\{1,1,2\}$

Notation : $\{1, 1, 2, 3\} = 1^2 2^1 3^1$

 $\mathsf{Perm}(1^2 2^3 3^1) = 1 \cdot \mathsf{Perm}(1^1 2^3 3^1) \sqcup 2 \cdot \mathsf{Perm}(1^2 2^2 3^1) \sqcup 3 \cdot \mathsf{Perm}(1^2 2^3)$

Retenir

Énumération lexicographique des multi-permutations :

$$\operatorname{Perm}(X) = \bigsqcup_{i=1}^{n} x_i \cdot \operatorname{Perm}(X/\{x_i\})$$



Coefficient multinomiaux :

L'union disjointe précédente nous donne l'égalité de cardinaux :

$$\binom{|I|}{i_1, i_2, \dots, i_k} = \binom{|I| - 1}{i_1 - 1, i_2, \dots, i_k} + \binom{|I| - 1}{i_1, i_2 - 1, \dots, i_k} + \cdots + \binom{|I| - 1}{i_1, i_2, \dots, i_k - 1}$$

On montre alors

$$\mathsf{Perm}(x_1^{i_1} \dots x_k^{i_k}). \, \mathsf{count}() = \frac{(i_1 + i_2 + \dots + i_k)!}{i_1! \, i_2! \, \dots \, i_k!} = \binom{|I|}{i_1, i_2, \dots, i_k}$$



Coefficient multinomiaux :

L'union disjointe précédente nous donne l'égalité de cardinaux :

$$\binom{|I|}{i_1, i_2, \dots, i_k} = \binom{|I| - 1}{i_1 - 1, i_2, \dots, i_k} + \binom{|I| - 1}{i_1, i_2 - 1, \dots, i_k} + \cdots + \binom{|I| - 1}{i_1, i_2, \dots, i_k - 1}$$

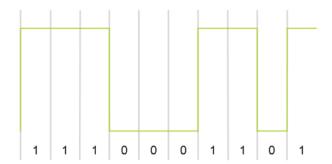
On montre alors :

$$\mathsf{Perm}(x_1^{i_1} \dots x_k^{i_k}). \, \mathsf{count}() = \frac{(i_1 + i_2 + \dots + i_k)!}{i_1! \, i_2! \, \dots \, i_k!} = \binom{|I|}{i_1, \, i_2, \, \dots, \, i_k}$$

Union disjointe 16 de 35

Autre application : transmission en codage NRZ

Non Return to Zero, Manière très élémentaire pour transmettre de l'information sur un ligne : 0 : -V, 1 : +V



Source: https://fr.wikipedia.org/wiki/Non_Return_to_Zero

Union disjointe 17 de 35

Perte de synchronisation en codage NRZ

Si on envoie une suite trop longue de bits identiques, le récepteur ne voit un signal constant. Il ne peut plus se baser sur les transitions pour se synchroniser et risque de perdre la synchronisation avec l'émetteur.

Definition

Une séquence de longueur n est dite non-repétitive d'ordre k (abréviation $NR_k(n)$) si elle ne contient pas se séquence de plus de k bits identiques consécutifs.

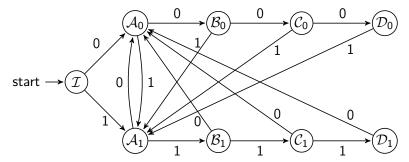
On note $NR_k(n, c, i)$ les suites qui commencent par au plus c i:

- $\blacksquare NR_k(n,0,0) = 1 \cdot NR_k(n-1,k-1,1)$
- $\blacksquare NR_k(n,c,0) = 0 \cdot NR_k(n-1,c-1,0) \sqcup 1 \cdot NR_k(n-1,k-1,1)$
- Idem en échangeant les rôles de 0 et 1.



disjointe 18 de 35

C'est un automate fini!



Tous les états sont acceptants.



Union récursive et automates finis

Retenir

La méthode précédente fonctionne pour touts les automates finis déterministes. Soit $\mathsf{Lang}_n(E)$ l'ensemble des mots de longueur n acceptés à partir de l'état E, alors on a la définition récursive :

■ Cas de base, pour un état E :

$$\mathsf{Lang}_0(E) = \begin{cases} \{\varepsilon\} & \textit{si E est terminal} \\ \emptyset & \textit{sinon} \end{cases}$$

■ Étape de la récursion :

$$\mathsf{Lang}_n(E) = \bigsqcup_{E \to E'} a \cdot \mathsf{Lang}_{n-1}(E')$$

Le produit cartesien 20 de 35

Le produit cartésien

Definition

On appelle **produit cartesien** de A et B l'ensemble C noté $C := A \times B$ défini par

$$C := \{(a, b) \mid a \in A, b \in B\}$$
.

- \blacksquare count(C) = count(A) · count(B) = $k \cdot \ell$
- On peut prendre la liste dans l'ordre lexicographique :



Le produit cartésien

Definition

On appelle **produit cartesien** de A et B l'ensemble C noté $C := A \times B$ défini par

$$C := \{(a, b) \mid a \in A, b \in B\}$$
.

- \blacksquare count(C) = count(A) · count(B) = $k \cdot \ell$
- On peut prendre la liste dans l'ordre lexicographique :





Itération sur un produit cartésien

Ordre lexicographique:

Itération en Python :

```
1 def iter_cartprod(A, B):
2 for a in A:
3 for b in B:
4 yield (a, b)
```



first, next sur un produit cartésien

```
def first_cartprod(A, B):
    return (A.first(), B.first())

def next_cartprod(A, B, x):
    (a , b) = x  # pattern matching
    try:
    return (a, B.next(b))
    except StopIteration:
    return (A.next(a), B.first())
```

rank sur un produit cartésien

```
def rank_cartprod(A, B, x):
    (a , b) = x  # pattern matching
    A.rank(a)*B.count() + B.rank(b)

def unrank_cartprod(A, B, i):
    c = B.count()
    return (A.unrank(i // c), B.unrank(i % c))
```



Notion de classe combinatoire

Définition (Classe combinatoire)

On appelle classe combinatoire un ensemble $\mathcal C$ dont les éléments e ont une taille (nommée aussi degrée) noté |e| et tels que l'ensemble $\mathcal C_n$ des éléments de taille n est fini :

$$\operatorname{count}(\{e \in \mathcal{C} \mid |e| = n\}) < \infty$$

Exemple:

- Les mots sur un alphabet où la taille est la longueur
- Les permutations de $\{1, \ldots n\}$ (taille = n)
- Les arbres binaires où la taille est le nombre de noeuds



L'union disjointe graduée

Si $\mathcal{C}=\mathcal{A}\sqcup\mathcal{B}$, les élements de \mathcal{A} et \mathcal{B} gardent leur taille dans l'union disjointe graduée :

$$C_n := A_n \sqcup B_n$$

Alors:

- \blacksquare \mathcal{C} . count $(n) = \mathcal{A}$. count $(n) + \mathcal{B}$. count(n)
- On peut prendre : C. list(n) = concat(A. list(n), B. list(n))
- → On peut réutiliser tout ce que l'on a vu sur les unions disjointes.



L'union disjointe graduée

Si $\mathcal{C}=\mathcal{A}\sqcup\mathcal{B}$, les élements de \mathcal{A} et \mathcal{B} gardent leur taille dans l'union disjointe graduée :

$$C_n := A_n \sqcup B_n$$

Alors:

- \blacksquare \mathcal{C} . count $(n) = \mathcal{A}$. count $(n) + \mathcal{B}$. count(n)
- On peut prendre : C. list(n) = concat(A. list(n), B. list(n))
- ⇒ On peut réutiliser tout ce que l'on a vu sur les unions disjointes.



Le produit cartesien gradué

Idée : les tailles (complexité, coût, nbr d'emplacements mémoires) s'ajoutent.

Definition

La taille de la paire $(a, b) \in \mathcal{A} \times \mathcal{B}$ est la somme des tailles :

$$|(a,b)|_{\mathcal{A}\times\mathcal{B}}:=|a|_{\mathcal{A}}+|b|_{\mathcal{B}}$$



Le produit cartesien gradué

Retenir

$$Si C = A \times B \ alors$$

$$C_n = \bigsqcup_{i+i-n} A_i \times B_j$$

Calcul de la cardinalité :

$$|\mathcal{C}_n| = \sum_{i+i=n} |\mathcal{A}_i| \times |\mathcal{B}_j| = \sum_{i=0}^n |\mathcal{A}_i| \times |\mathcal{B}_{n-i}|$$

On peut alors prendre l'ordre union/lexicographique suivant

$$A_0 \times B_n \mid A_1 \times B_{n-1} \mid A_2 \times B_{n-2} \mid \dots \mid A_n \times B_0$$



Le produit cartesien gradué

Retenir

$$Si C = A \times B \ alors$$

$$C_n = \bigsqcup_{i+i=n} A_i \times B_j$$

Calcul de la cardinalité :

$$|\mathcal{C}_n| = \sum_{i+j=n} |\mathcal{A}_i| \times |\mathcal{B}_j| = \sum_{i=0}^n |\mathcal{A}_i| \times |\mathcal{B}_{n-i}|$$

On peut alors prendre l'ordre union/lexicographique suivant :

$$A_0 \times \mathcal{B}_n \mid A_1 \times \mathcal{B}_{n-1} \mid A_2 \times \mathcal{B}_{n-2} \mid \dots \mid A_n \times \mathcal{B}_0$$



Application les arbres binaires

Spécification récursive :

$$\mathcal{B}in\mathcal{T}ree = Leaf \sqcup Node(\mathcal{B}in\mathcal{T}ree \times \mathcal{B}in\mathcal{T}ree)$$

Deux manières de compter les tailles :

1 Nombre de feuille :

$$\mathcal{B}\mathsf{in}\mathcal{T}\mathsf{ree} = \mathsf{Leaf}_1 \sqcup \mathcal{B}\mathsf{in}\mathcal{T}\mathsf{ree} \times \mathcal{B}\mathsf{in}\mathcal{T}\mathsf{ree}$$

2 Nombre de Noeuds :

```
\mathcal{B}in\mathcal{T}ree = Leaf_0 \sqcup Node_1 \times \mathcal{B}in\mathcal{T}ree \times \mathcal{B}in\mathcal{T}ree
```



Application les arbres binaires

Spécification récursive :

$$\mathcal{B}in\mathcal{T}ree = Leaf \sqcup Node(\mathcal{B}in\mathcal{T}ree \times \mathcal{B}in\mathcal{T}ree)$$

Deux manières de compter les tailles :

1 Nombre de feuille :

$$\mathcal{B}\mathsf{in}\mathcal{T}\mathsf{ree} = \mathsf{Leaf}_1 \ \sqcup \ \mathcal{B}\mathsf{in}\mathcal{T}\mathsf{ree} \times \mathcal{B}\mathsf{in}\mathcal{T}\mathsf{ree}$$

2 Nombre de Noeuds :

```
\mathcal{B}\mathsf{in}\mathcal{T}\mathsf{ree} = \mathsf{Leaf}_0 \sqcup \mathsf{Node}_1 \times \mathcal{B}\mathsf{in}\mathcal{T}\mathsf{ree} \times \mathcal{B}\mathsf{in}\mathcal{T}\mathsf{ree}
```



Liste de tous les arbres à *n* Nœuds

Algorithme

```
■ Entrée : un entier positif ou nul n
```

■ **Sortie** : une liste d'arbres

Nombre de Catalan

Proposition

Le nombre d'arbres binaires à n nœuds est appelé n-ième nombre de Catalan noté C_n . Les nombre de Catalan vérifient la récurrence :

$$C_0 = 1$$
 $C_n = \sum_{i=0}^{n-1} C_i C_{n-1-i}$.

On peut trouver une formule close :

$$C_n=\frac{(2n)!}{n!(n+1)!}.$$

Voici les premières valeurs :

$$C_0 = 1$$
, $C_1 = 1$, $C_2 = 2$, $C_3 = 5$, $C_4 = 14$, $C_5 = 42$, $C_6 = 132$.



Specification d'une classe combinatoire

But : on veut décrire une classe combinatoire de manière à pouvoir appliquer automatiquement les algorithmes de comptage, itération, génération,....

Pour ceci, on va utiliser une **grammaire** qui code comment appliquer récursivement les constructions précédentes.



Specification d'une classe combinatoire

Retenir (Grammaire de description d'objets)

Constructeurs terminaux :

- lacktriangledown $\mathcal{E}(o)$: la classe qui contient un seul objet o de taille 0
- lacksquare $\mathcal{Z}(o)$: la classe qui contient un seul objet o de taille 1

Constructeurs binaires :

■ C = Union(A, B) : union disjointe graduée :

$$|a|_{\mathcal{C}} = |a|_{\mathcal{A}} \text{ si } a \in \mathcal{A} \qquad |a|_{\mathcal{C}} = |b|_{\mathcal{B}} \text{ si } b \in \mathcal{B}$$

■ C = Prod(A, B): produit cartésien graduée :

$$|(a,b)|_{\mathcal{C}} = |a|_{\mathcal{A}} + |b|_{\mathcal{B}} \text{ si } a \in \mathcal{A} \text{ et } b \in \mathcal{B}$$

Exemples:

■ Les arbres binaires, la taille est le nombres de Noeuds

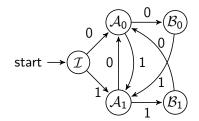
$$\mathcal{B}in\mathcal{T}ree = Union(\mathcal{E}(\bot), Prod(\mathcal{Z}(\circ), Prod(\mathcal{B}in\mathcal{T}ree, \mathcal{B}in\mathcal{T}ree))$$

■ Les arbres binaires, la taille est le nombres de Feuilles

$$\mathcal{B}in\mathcal{T}ree = Union(\mathcal{Z}(\bot), Prod(\mathcal{B}in\mathcal{T}ree, \mathcal{B}in\mathcal{T}ree))$$



Exemples: Codage d'un automate fini

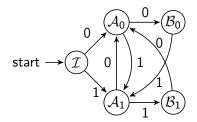


Les mots binaires qui n'ont pas plus de deux lettres identiques consécutives :

$$\begin{split} \mathcal{I} &= \mathcal{E}(\varepsilon) \; \sqcup \; \left(\mathcal{Z}("0") \times \mathcal{A}_{0}\right) \; \sqcup \; \left(\mathcal{Z}("1") \times \mathcal{A}_{1}\right) \\ \mathcal{A}_{0} &= \mathcal{E}(\varepsilon) \; \sqcup \; \left(\mathcal{Z}("0") \times \mathcal{B}_{0}\right) \; \sqcup \; \left(\mathcal{Z}("1") \times \mathcal{A}_{1}\right) \\ \mathcal{B}_{0} &= \mathcal{E}(\varepsilon) \; \sqcup \; \left(\mathcal{Z}("1") \times \mathcal{A}_{1}\right) \\ \mathcal{A}_{1} &= \mathcal{E}(\varepsilon) \; \sqcup \; \left(\mathcal{Z}("0") \times \mathcal{A}_{0}\right) \; \sqcup \; \left(\mathcal{Z}("1") \times \mathcal{B}_{1}\right) \\ \mathcal{B}_{1} &= \mathcal{E}(\varepsilon) \; \sqcup \; \left(\mathcal{Z}("0") \times \mathcal{A}_{0}\right) \end{split}$$



Exemples: Codage d'un automate fini



Dans cet exemple : plus simple, on inclut la lettre dans l'état :

$$\mathcal{I} = \mathcal{E}(\varepsilon) \; \sqcup \; \mathcal{A}_0 \; \sqcup \; \mathcal{A}_1$$

$$\mathcal{A}_0 = \mathcal{Z}("0") \times (\mathcal{E}(\varepsilon) \; \sqcup \; \mathcal{B}_0 \sqcup \; \mathcal{A}_1)$$

$$\mathcal{B}_0 = \mathcal{Z}("0") \times (\mathcal{E}(\varepsilon) \; \sqcup \; \mathcal{A}_1)$$

$$\mathcal{A}_1 = \mathcal{Z}("1") \times (\mathcal{E}(\varepsilon) \; \sqcup \; \mathcal{B}_1 \sqcup \; \mathcal{A}_0)$$

$$\mathcal{B}_1 = \mathcal{Z}("1") \times (\mathcal{E}(\varepsilon) \; \sqcup \; \mathcal{A}_0)$$