Journée Optimisation dans les Réseaux
19 juin 2015

Problèmes de graphes et protection de la biodiversité : quelques exemples

Alain Billionnet

Développement durable

« Développement qui répond aux besoins du présent sans compromettre la capacité des générations futures à répondre aux leurs »

Rapport de la commission des nations unies sur l’environnement et le développement (1987)

Finalités du développement durable

- **Lutte contre le changement climatique**
- **Préservation de la biodiversité**
- **Cohésion sociale et solidarité entre les territoires et les générations**
- **Épanouissement de tous les êtres humains**
- **Développement suivant des modes de production et de consommation responsables**

Biodiversité

(JORF 2009)

Diversité des organismes vivants, qui s'apprécie en considérant la diversité des espèces, celle des gènes au sein de chaque espèce, ainsi que l'organisation et la répartition des écosystèmes

**Stopper la perte de biodiversité :
un des grands défis auxquels la communauté internationale est confrontée**

Modèles et méthodes d’optimisation

$$\downarrow$$

Outils d’aide à la décision pour la protection de la biodiversité

- **Réserves naturelles**
- **Fragmentation des paysages**
- **Diversité phylogénétique**

I

Réserves naturelles

- **De nombreux pays ont élaboré des stratégies pour stopper la perte de biodiversité**

- **Aires protégées (ou réserves naturelles) : contribution décisive à ces stratégies**

Protection directe des éléments de la biodiversité risquant de disparaître

→ Avenir garanti pour chaque espèce ou habitat menacé

Ressources limitées

pour protéger ou restaurer les zones concernées

investir de façon efficace

I.1

Sélection de réserves naturelles : problème de base et variantes

Données

- **Parcelles (patchs, taches, ilôts) : $\mathbf{P}_1, \mathbf{P}_2, \ldots, \mathbf{P}_n$**
- **Espèces à protéger : $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m$**
- **On connaît S_k : parcelles permettant de protéger $\mathbf{e}_k$ (une parcelle de S_k sélectionnée $\Rightarrow$ $\mathbf{e}_k$ protégée)**

Problème

Sélectionner un nombre minimal de parcelles permettant de protéger toutes les espèces considérées

Formulation par un programme linéaire en 0-1

(problème de recouvrement)

$x_i = 1 \Leftrightarrow$ parcelle P_i sélectionnée

$$\begin{cases} \min \sum_{i=1}^{n} x_i \\ \text{s.c.} \left| \begin{array}{l} \sum_{i \in S_k} x_i \geq 1 \quad \forall \text{espèce } e_k \\ x \in \{0,1\}^n \end{array} \right. \end{cases}$$

S_k : parcelles contenant e_k

Autre objectif : $\min \sum_{i=1}^{n} c_i x_i$

c_i : coût associé à P_i

Une variante : richesse spécifique

Maximiser le <u>nombre</u> d'espèces protégées

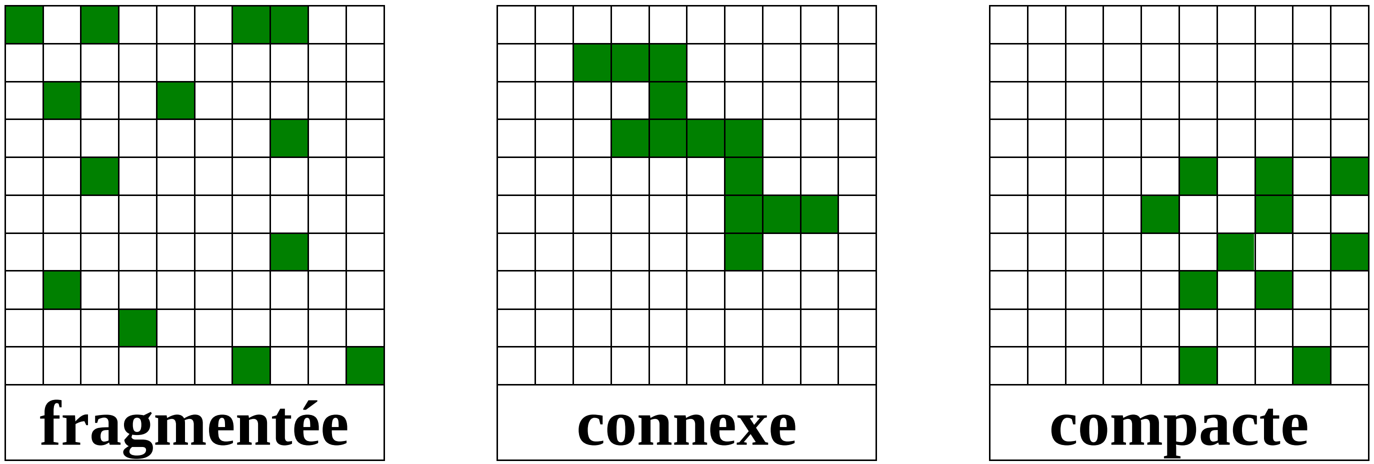
$$\begin{cases} \max \sum_{k=1}^{m} y_k \\ \text{s.c.} \left| \begin{array}{ll} y_k \le \sum_{i \in S_k} x_i & \forall \text{ espèce } e_k \\ \sum_{i=1}^{n} c_i x_i \le B & \\ x_i \in \{0,1\},\ y_k \in \{0,1\} & \end{array} \right. \end{cases}$$

$y_k = 1$ ssi au moins une des parcelles retenues protége e_k

Contrainte budgétaire

Nombreuses variantes du problème dans la littérature de la conservation biologique :

Nombre ou coût des <u>parcelles</u> protégées, nombre d'<u>espèces</u> protégées, <u>connexité</u> des parcelles retenues, <u>compacité</u> des parcelles retenues, aspect <u>probabiliste</u>, …



I.II

Sélection d'une réserve connexe et compacte

Connexité et compacité :

Augmentation des chances de survies des espèces

- Déplacements facilités
  (migration, fuite, nourriture)

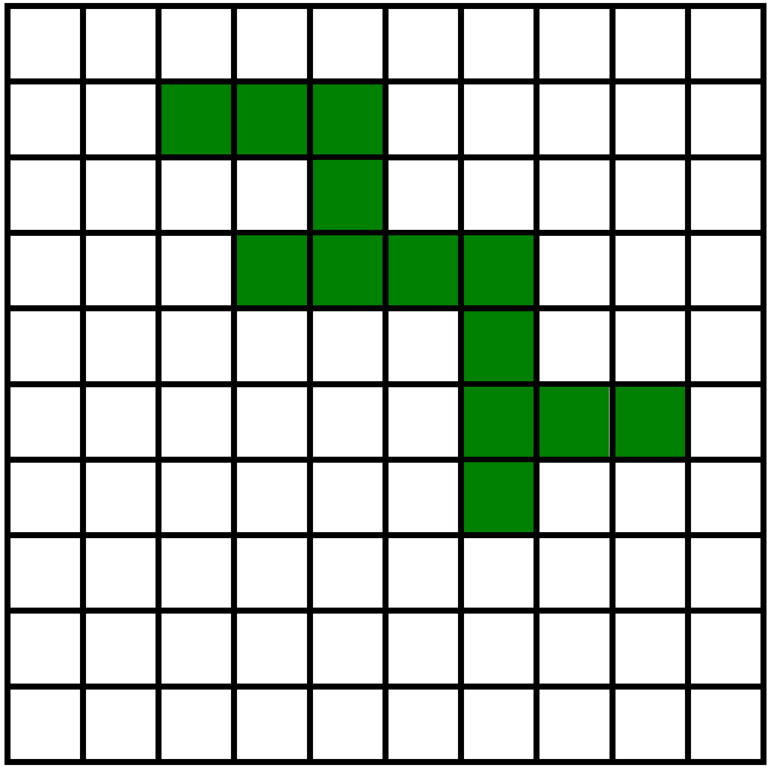
- Augmentation de la transmission des gènes

- Diminution des effets de bordure
  (détérioration de l'habitat)

- **Espèces :** $e_1, e_2, \ldots, e_m$
- **Zone étudiée : matrice de $m \times n$ parcelles carrées**
- **Effectif de l'espèce e_k dans la parcelle P_{ij} : a_{ijk}**
- **Une espèce e_k est protégée si sa population dans la réserve est supérieure ou égale à b_k**
- **Parcelles adjacentes $\Leftrightarrow$ un côté commun**

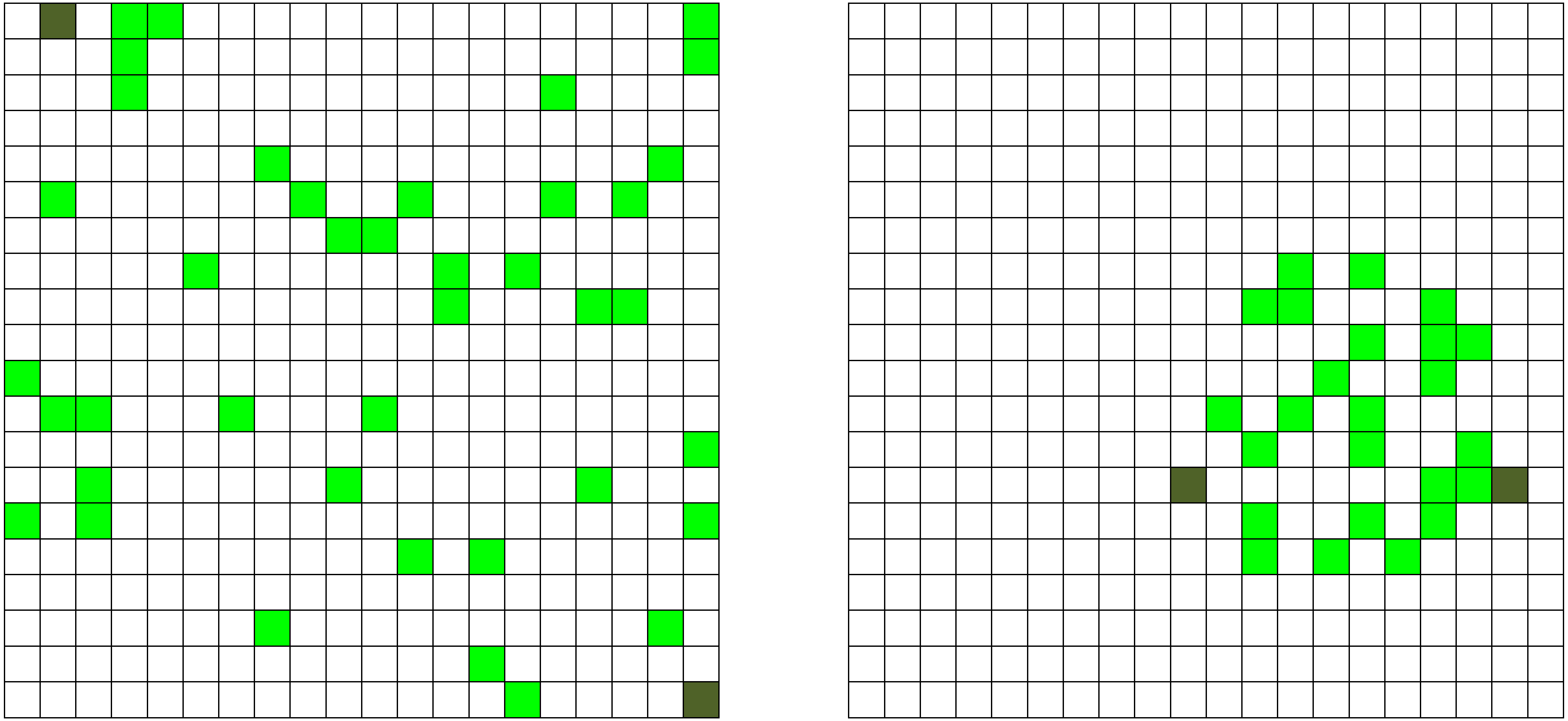
Connexité

Mouvement possible entre 2 parcelles quelconques de la réserve sans quitter la réserve

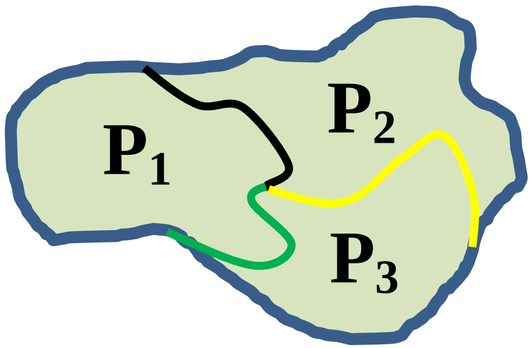


Mesures possibles de la compacité

- **Distance entre les 2 parcelles les plus éloignées**



- $\dfrac{\textbf{Périmètre total de la réserve}}{\textbf{Aire totale de la réserve}}$



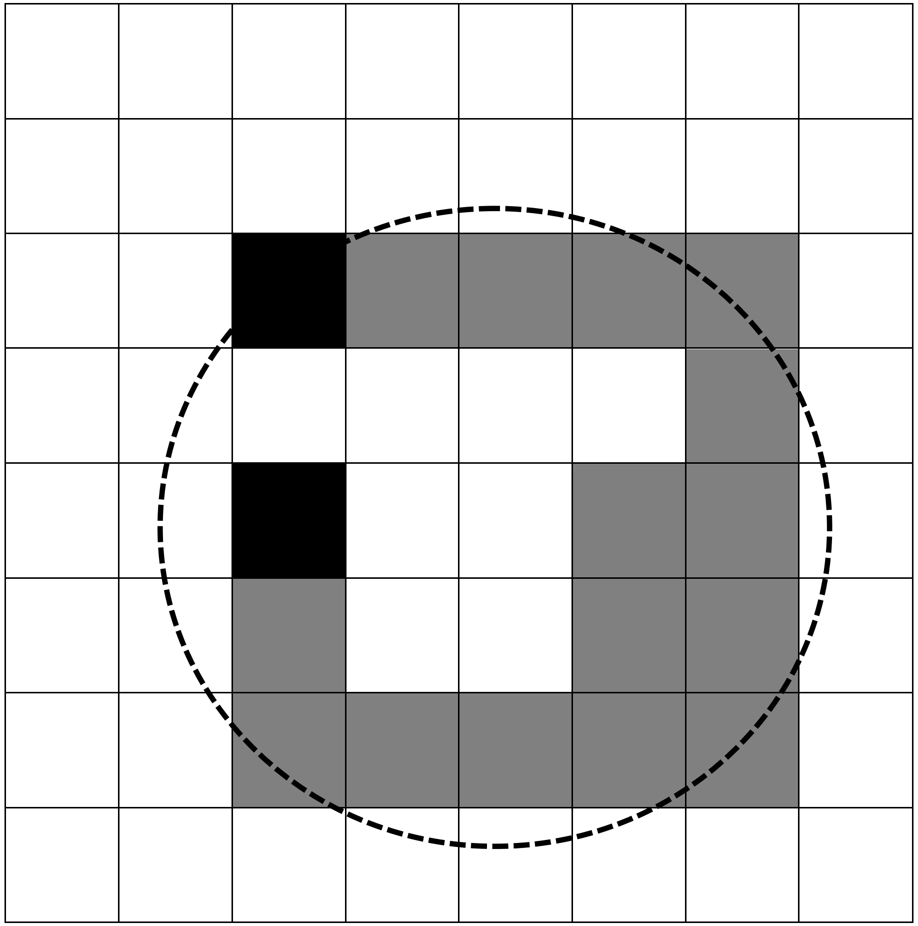
longueur frontière commune P_i / P_j

$$\frac{\sum_{i=1}^{n} l_i x_i - 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} l_{ij} x_i x_j}{\sum_{i=1}^{n} a_i x_i}$$

périmètre de P_i

aire de P_i

- **Rayon du plus petit cercle contenant la réserve**



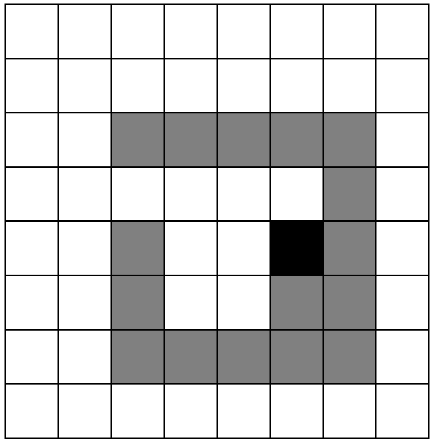
Critère retenu

- d_{ijkl} : distance minimale à parcourir pour relier P_{ij} et P_{kl} sans quitter la réserve R

- $\forall$ parcelle de R : $$e(P_{ij}) = \max\left\{d_{ijkl} : P_{kl} \in R\right\}$$

- Compacité :

$$\min\left\{e(P_{ij}) : P_{ij} \in R\right\}$$



$$e(P_{56}) = 7$$

Problème : trouver un ensemble de parcelles qui protègent un nombre maximal d'espèces et qui respectent les contraintes :

- **de budget**

- **de connexité**

- **de compacité :** $\min\left\{ e(P_{ij}) : P_{ij} \in \text{Réserve} \right\} \leq r$

Un graphe connexe $G = (X, E)$

- **$d(x, y)$ longueur de la chaîne de longueur minimale reliant x et y**

- **Écartement de x : $e(x) = \max\{d(x, y) : y \in X\}$**

- **Centre de G : sommet x_0 d'écartement minimal**

- **Rayon de G : écartement de x_0**

Formulation en termes de graphes

Ensemble des parcelles $\rightarrow$ graphe $G = (X, E)$

- **Sommets : parcelles**
- **Arêtes : $\{P_{ij}, P_{kl}\}$ ssi P_{ij} et P_{kl} sont adjacentes**

1,1	1,2	1,3	1,4
2,1	2,2	2,3	2,4
3,1	3,2	3,3	3,4
4,1	4,2	4,3	4,4

Problème : trouver $S \subset X$ tel que

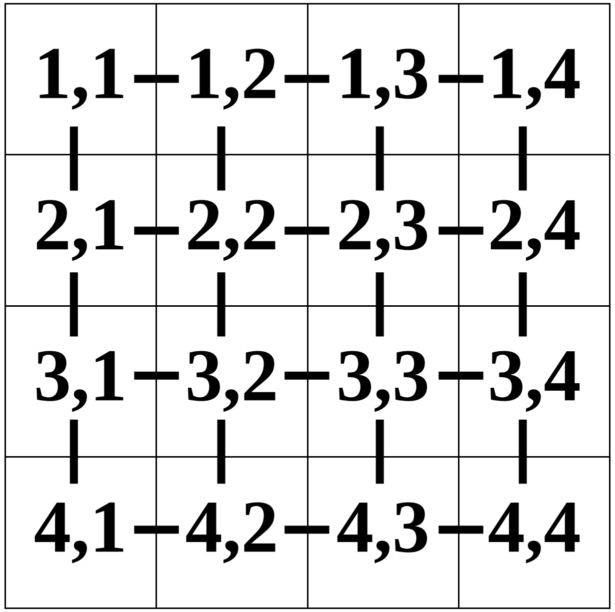
- sous-graphe induit par S connexe

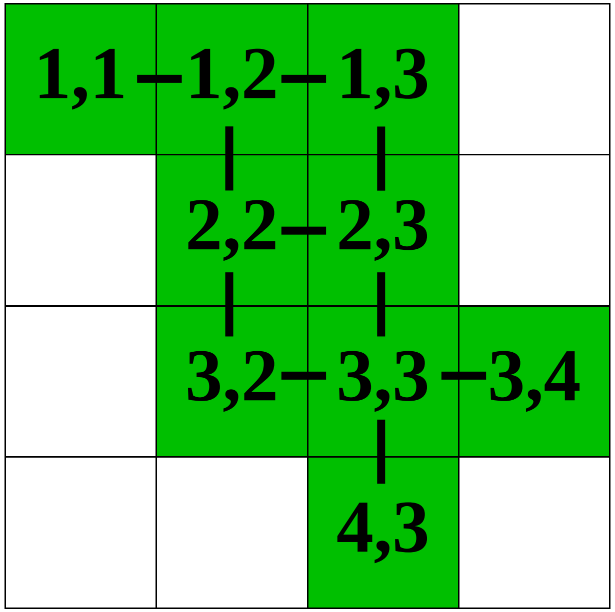
- de coût $\leq B$

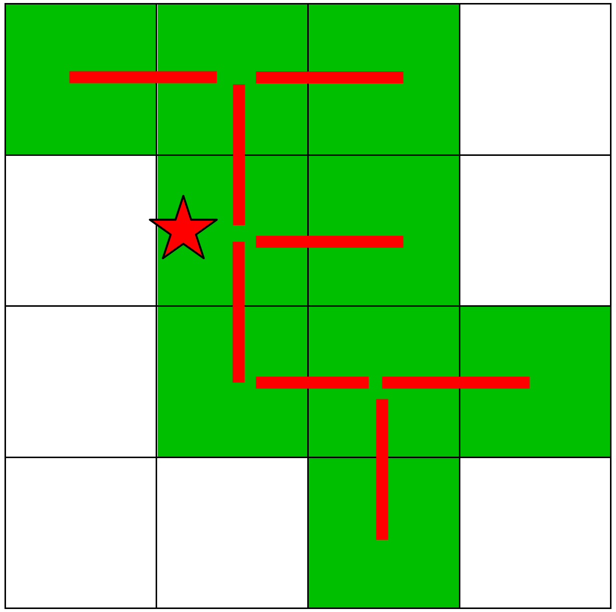
- qui protège le plus grand nombre d'espèces

- de rayon $\leq r$

$\Leftrightarrow$ qui admet un arbre couvrant de hauteur $\leq r$


1,1 1,2 1,3 1,4
2,1 2,2 2,3 2,4
3,1 3,2 3,3 3,4
4,1 4,2 4,3 4,4

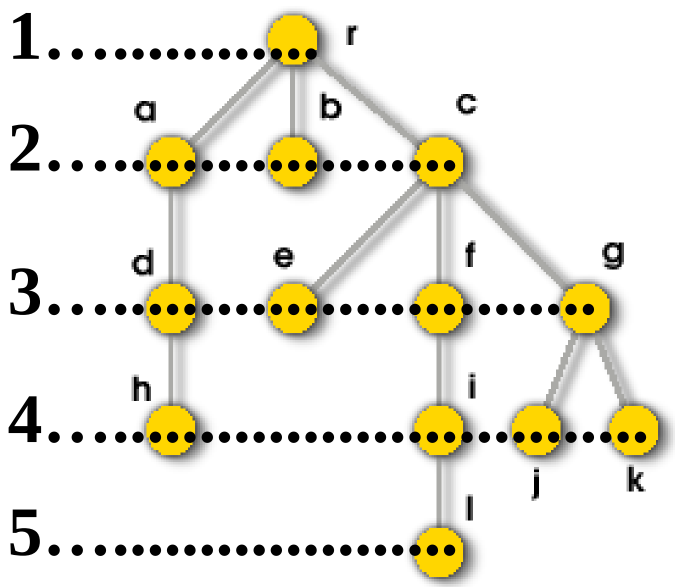

1,1 1,2 1,3
2,2 2,3
3,2 3,3 3,4
4,3



Modélisation par un PL en variables 0-1

Variables booléennes

- $x_{ijh} = 1$ ssi parcelle P_{ij} retenue au niveau h de l'arbre ($\Rightarrow \exists$ chaîne de longueur $\leq h-1$ entre la racine et P_{ij})



$$\begin{cases} \max \displaystyle\sum_{k \text{ espèce}} y_k \\ \text{s.c.} \left| \begin{array}{l} \displaystyle\sum_{(i,j)\text{parcelle}} c_{ij} w_{ij} \le B \\ w_{ij} = \sum_{h=1}^{r+1} x_{ijh} \qquad \forall \text{ parcelle } (i,j) \\ \displaystyle\sum_{(i,j)\text{parcelle}} a_{ijk} w_{ij} \ge b_k\, y_k \qquad \forall \text{ espèce } k \\ \text{Cont. de connexité et compacité (rayon} \le r) \\ w_{ij}, x_{ijh}, y_k \in \{0,1\} \end{array} \right. \end{cases}$$

- $w_{ij} = 1$ ssi P_{ij} retenue
- $y_k = 1$ ssi e_k protégée

Connexité et compacité

Une parcelle est choisie comme racine

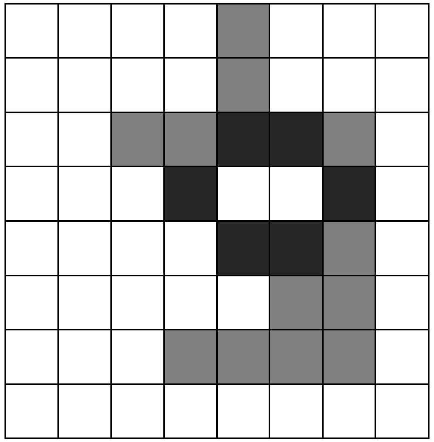
$$\sum_{P_{ij}\ \text{parcelle}} x_{ij1} = 1$$

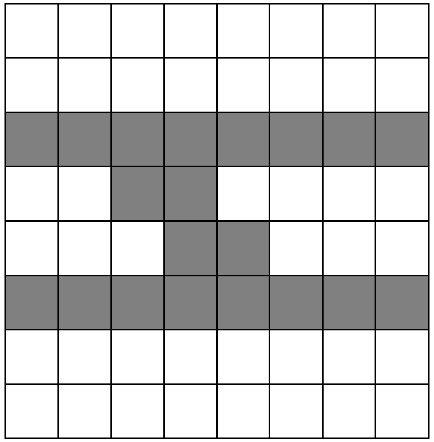
$$x_{ijh} \le \sum_{P_{kl}\ \text{adjacente à}\ P_{ij}} x_{k,l,h-1} \qquad \forall\ \text{parcelle}\ P_{ij}, \forall h \ge 2$$

P_{ij} affectée au niveau *h*

$\Rightarrow$ au moins une parcelle adjacente affectée au niveau *h*-1

Connexité des parcelles extérieures à la reserve





$$\sum_{P_{ij} \in E_2} x_{ij} \geq |E_2| \times (1 - \sum_{P_{ij} \in E_1} (1 - x_{ij}))$$

E_1 : ens. de parcelles « isolantes »

E_2 : ens. de parcelles « isolées » par les parcelles de E_1

Conclusion

- **Approche efficace pour 100 espèces et 400 parcelles**
- **Connexité : problème classique (difficile) de graphes**
- **Définition réaliste de la compacité ne compliquant pas la formulation de la connexité**
- **Prise en compte de la connexité à l'extérieur de la réserve**
- **Une façon de traiter le problème**
- **Parcelles candidates : pas obligatoirement une grille**

II

Fragmentation des paysages

Corridors biologiques

Développement urbain et agricole
exploitation des forêts

⇓

Fragmentation des paysages

⇓

Les espèces ne peuvent se déplacer
comme elles le devraient

⇓

Diminution de la biodiversité

Matw (Wikipedia allemand)

Massifs boisés non connectés
Nombreux invertébrés isolés dans les boisements
Sangliers et cervidés peuvent encore circuler

Rufus46 (GFDL)

Boisements relativement connectés

Ligne TGV avec une double clôture


Senet (personal work)

Tunnel et viaduc : beaucoup moins fragmentant

Corridors biologiques : la trame verte et bleue

Mesure phare du Grenelle Environnement

« Porte l'ambition d'enrayer le déclin de la biodiversité au travers de la préservation et de la restauration des continuités écologiques »

(Ministère de l'écologie, du développement durable et de l'énergie)

Les corridors biologiques

- « Chemins » empruntés par les espèces animales pour se déplacer, se reproduire, fuir, migrer, etc.

- Essentiels au développement des différentes espèces et donc à la biodiversité

- Dépendent des espèces

- Restauration des réseaux de corridors : une des grandes stratégies de protection des espèces menacées par la fragmentation de leur habitat

Bruno Robert (Œuvre personnelle)

Les anciennes voies ferrées peuvent avoir des fonctions de corridors jusque dans les zones urbaines et industrielles

Un premier problème :

Restauration optimale de parcelles pour connecter un ensemble de réserves

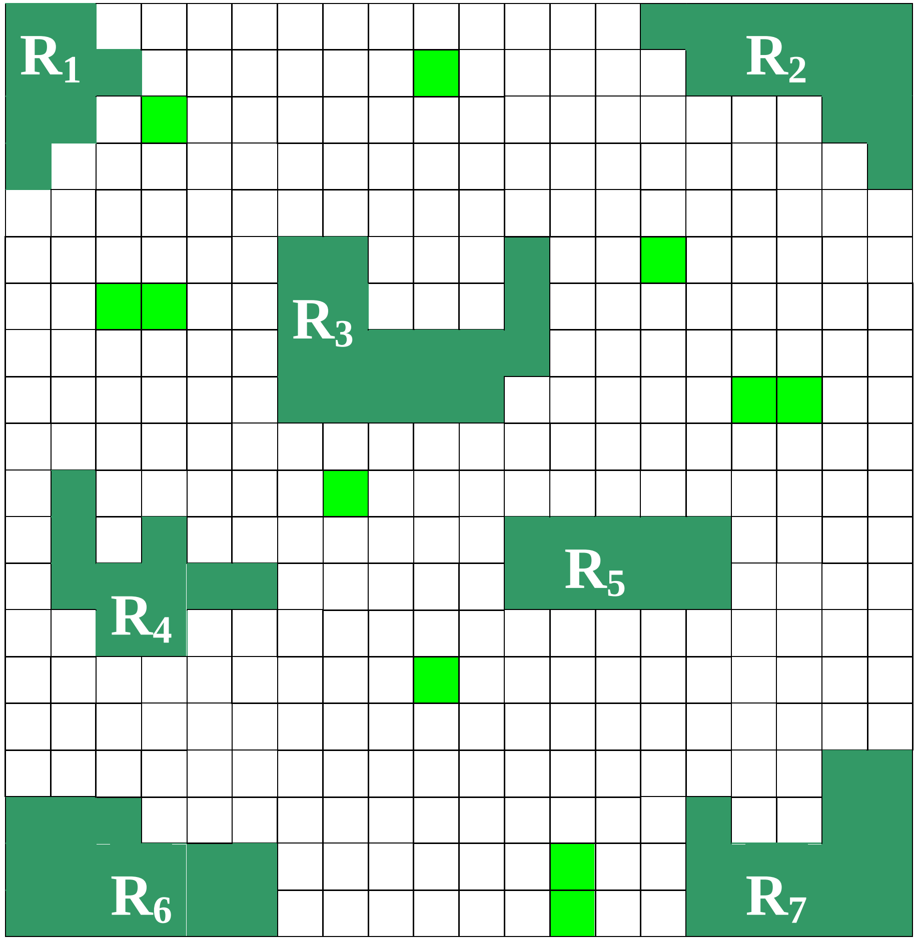
- **Zone étudiée : ensemble de parcelles disjointes**
- **Plusieurs états possibles pour chaque parcelle**

Irrémédiablement dégradée	**Hautement dégradée**	**Partiellement dégradée**	**Restaurée**

- **Ensemble de réserves (sous-ensemble connexe de parcelles) = zones d'habitat de certaines espèces**

- **Parcelles adjacentes : un côté commun**

Problème : Parcelles à restaurer pour relier entre elles toutes les réserves, au moindre coût ?



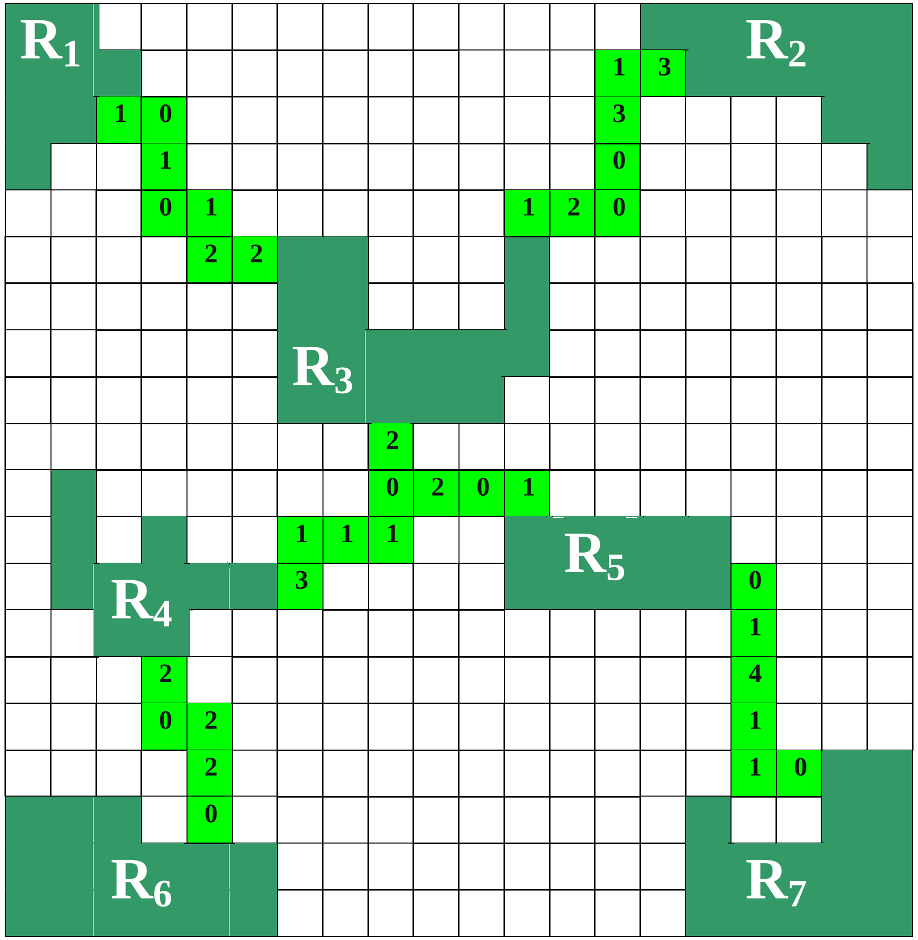
Quelles parcelles faut-il restaurer de façon à connecter les 7 réserves $R_1, R_2, \ldots, R_7$ au moindre coût ?

c_i : coût de restauration de la parcelle P_i

Parcelle habitat intégrée à une réserve

Parcelle habitat non intégrée à une réserve

Parcelle non habitat



Restauration de 34 parcelles ⟹ connexité

Coût : 41 unités - Longueur corridor $R_6 \rightarrow R_7$: 29 parcelles

Formulation en termes de graphes

Graphe associé $G = (X, U)$

X : **ensemble des parcelles**

U : $(P_i, P_j) \in U$ **ssi** P_i **et** P_j **sont adjacentes**

P_{i_k} : **parcelle représentant la réserve** R_k

Déterminer

Un ens. de sommets $S \subset X$ et un ens. d'arcs $A \subset U$ t.q.

- **$G = (S \bigcup_{k=1,\ldots,N} \mathrm{P}_{i_k}, A)$ admette un chemin de la parcelle représentant la réserve R_N à la parcelle représentant la réserve R_k ($k = 1,\ldots, N$-1)**
- **coût de restauration $\sum_{\mathrm{P}_i \in S} c_i$ minimal**

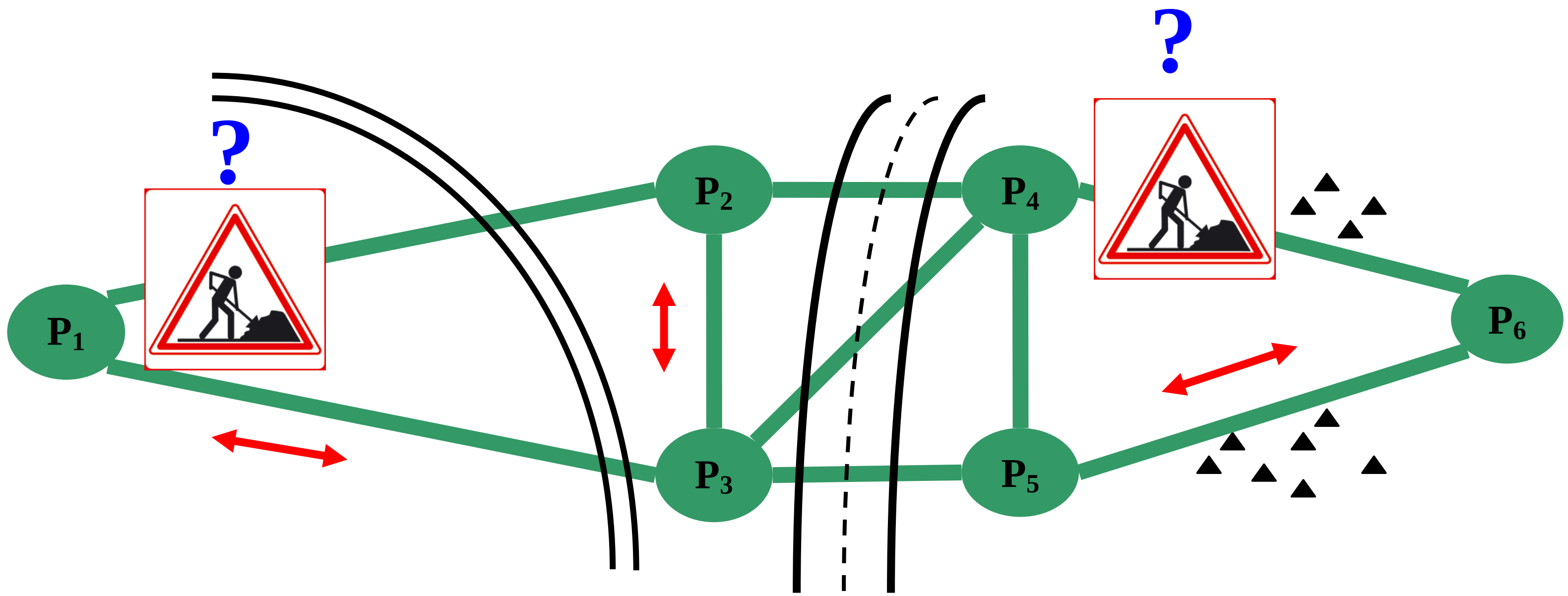
Arborescence de Steiner

Contrainte: limiter la longueur des corridors

Un deuxième problème

Maximiser la perméabilité d’un réseau de corridors

Parcelles : P_1, P_2,…, P_6 - Corridors : [P_1,P_2], [P_1,P_3],…,[P_5,P_6]



Objectif

Améliorer au maximum la perméabilité du réseau en aménageant les corridors sous une contrainte budgétaire

Perméabilité : espérance mathématique de la distance parcourue par les espèces considérées

Photo RFF

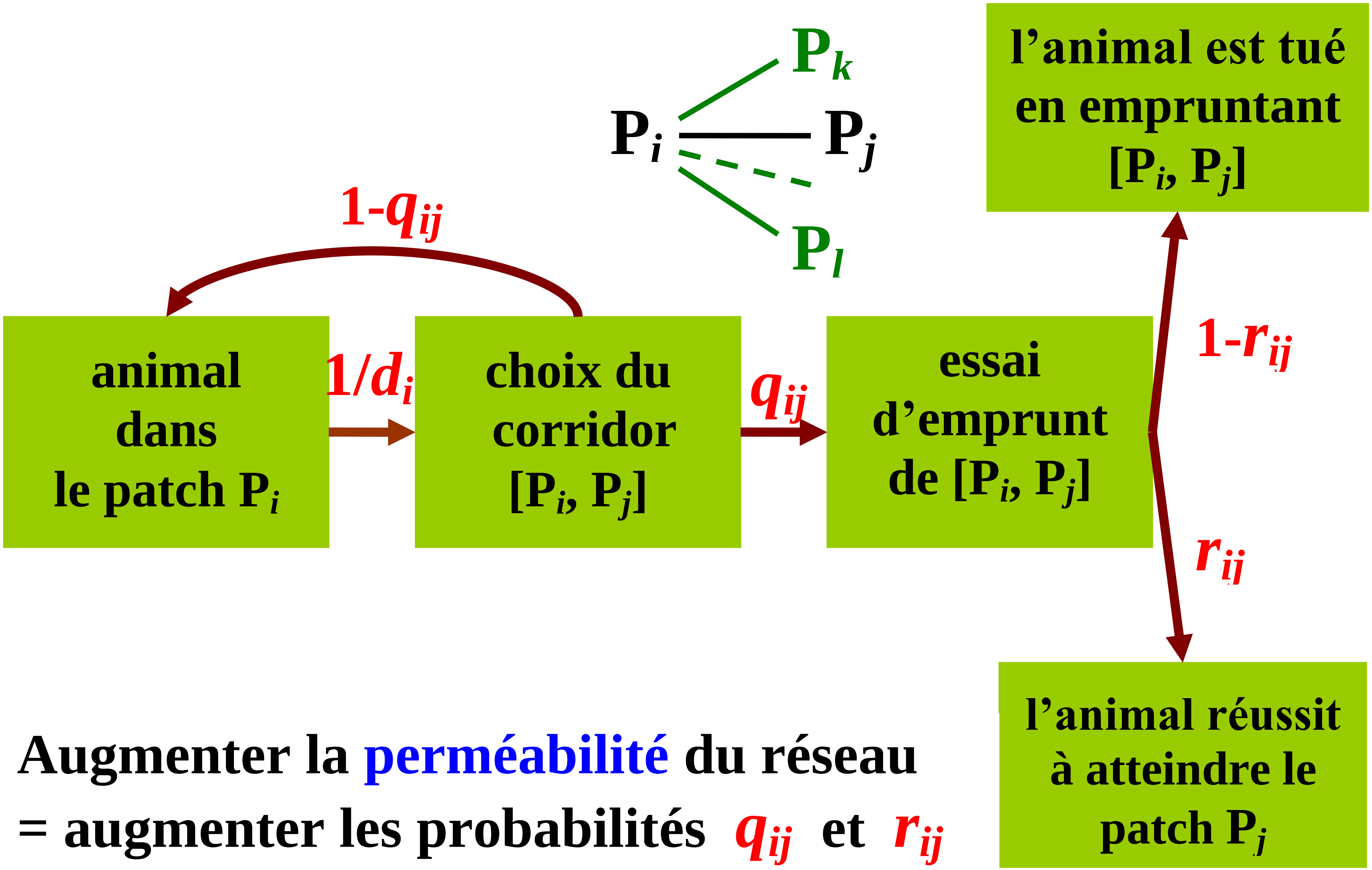
Passage grande faune sur une ligne TGV

CANOPÉ académie d'Amiens

Passage grande faune sur A26

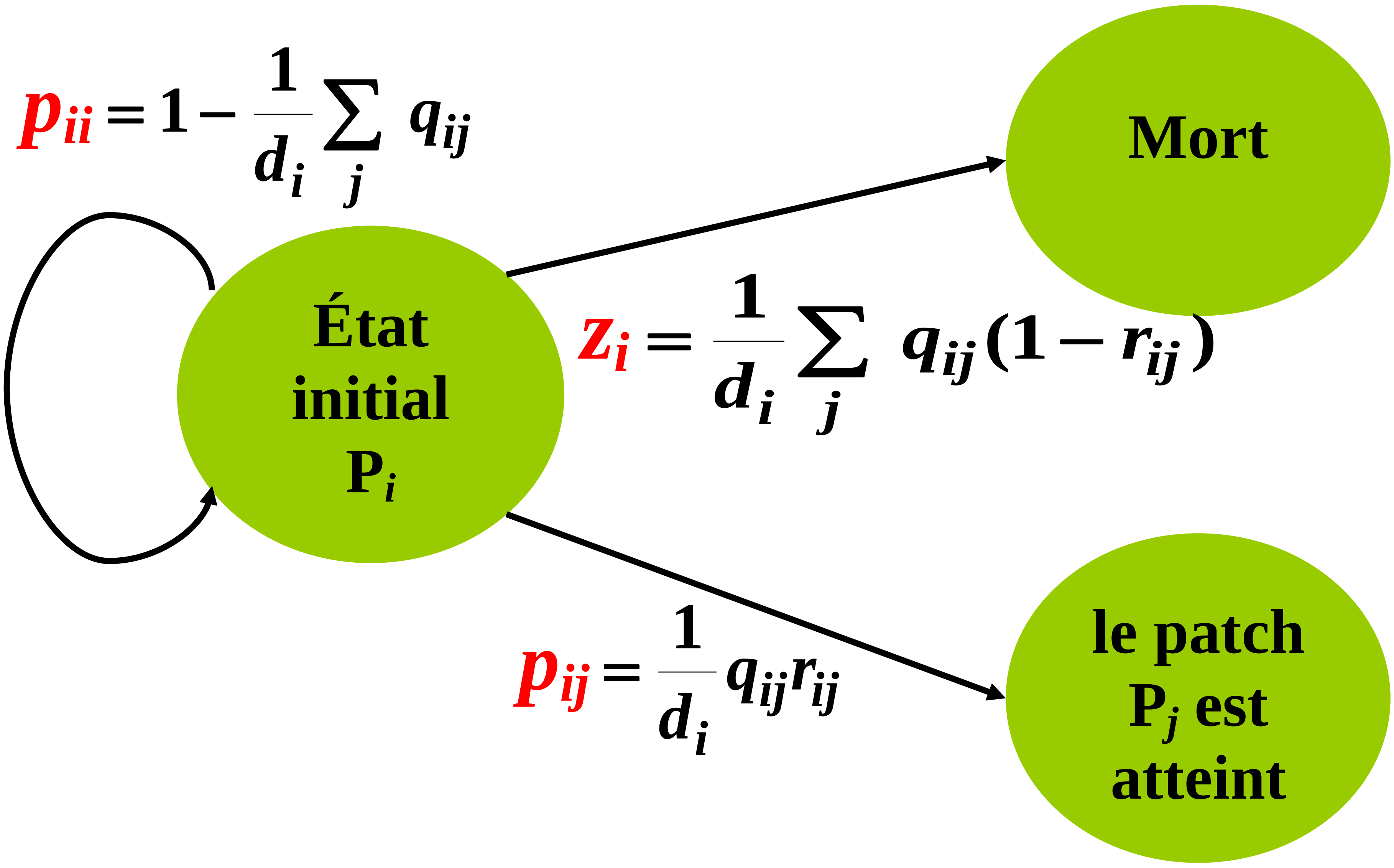
Recensement national des passages pour la grande faune : inventaire régions Champagne-Ardennes, Alsace Lorraine. Carsignol J., 1991.

Passage à chevreuil sous l'A4



Augmenter la perméabilité du réseau

= augmenter les probabilités q_{ij} et r_{ij}


$p_{ii} = 1 - \frac{1}{d_i} \sum_j q_{ij}$
Mort
État
initial
P_i
$z_i = \frac{1}{d_i} \sum_j q_{ij}(1 - r_{ij})$
$p_{ij} = \frac{1}{d_i} q_{ij} r_{ij}$
le patch
P_j est
atteint

Chaîne de Markov associée

	P_1	P_2	…	P_n	P_{n+1}
P_1	p_{11}	p_{12}	…	p_{1n}	$p_{1,n+1}$
P_2	p_{21}	p_{22}	…	p_{2n}	$p_{2,n+1}$
…	…	…	…	…	…
P_n	p_{n1}	p_{n2}	…	p_{nn}	$p_{n,n+1}$
P_{n+1}	0	0	0	0	1

n **états transitoires :** P_1 , P_2 , … , P_n

1 état absorbant : P_{n+1}

$$
\begin{array}{c|c|c|}
 & \textbf{Trans.} & \textbf{Abs.} \\
\hline
\textbf{Trans.} & Z & D \\
\hline
\textbf{Abs.} & 0 & I \\
\hline
\end{array}
$$

Matrice des probabilités de transition

Théorème : $N = (I - Z)^{-1}$

n_{ij} **= nb. moyen de passages par** j **partant de** i **avant abs.**

$\Rightarrow$ **nb. moyen de passages de** k **à** l **partant de** i **avant abs.**

$= n_{ik}\, p_{kl}$

Variables : p_{ij} **(probabilité de transition)**

Esp. math. de la distance totale parcourue
(n individus initialement dans $P_1, \dots, P_n$)

$$\begin{cases} \max\ f(\Pi) \\ \text{s.c.} \left| \ \Pi = \begin{pmatrix} Z & D \\ 0 & I \end{pmatrix} \in \tilde{\Pi} \subset M_{n+1} \right. \end{cases}$$

Ensemble des matrices
stochastiques $(n+1)\times(n+1)$

Variables

p_{ij} : probabilité de transition

w_i : $\sum_j (\text{nb. moyen de passages par } \mathrm{P}_i \text{ partant de } \mathrm{P}_j)$

longueur du corridor $[\mathrm{P}_i,\mathrm{P}_j]$

$$\begin{cases} \max \sum_{\text{corridor}[\mathrm{P}_i,\mathrm{P}_j]} l_{ij}\, \underbrace{(w_i p_{ij} + w_j p_{ji})} \\ \text{s.c.} \left| \begin{array}{l} w_i - \sum_{j=1}^{n} w_j p_{ji} = 1 \quad \forall \text{patch } i \\ \Pi \in \tilde{\Pi} \end{array} \right. \end{cases}$$

nombre de passages dans $[\mathrm{P}_i,\mathrm{P}_j]$

Difficultés

- **Exprimer p_{ij} en fonction des aménagements réalisés dans le corridor $[P_i , P_j]$**

- **non linéarité du programme : $w_i\, p_{ij}$**

$x_{ijk} = 1 \;\Leftrightarrow$ invest[t] n°k réalisé dans le corridor $[P_i , P_j]$

$c_{ijk} =$ coût investissement n°k dans le corridor $[P_i , P_j]$

$\Rightarrow$ **Programme quadratique en variables mixtes**

Conclusion

- **Une dizaine de patchs**
- **Une vingtaine de corridors**
- **≈ 5 niveaux d'aménagement différents**
- **Possibilité d'ajouter des contraintes supplémentaires par exemple sur le nombre de passages dans chaque corridor ⇒ augmentation du temps de calcul**
- **Grandes instances : heuristiques / chaînes de Markov (Travaux existants : heuristiques / simulations)**

III

Diversité phylogénétique

Arbre phylogénétique

$T = (Y, X, A)$

Y **= ensemble des sommets non feuilles**

X **= ensemble des sommets feuilles (espèces, taxons)**

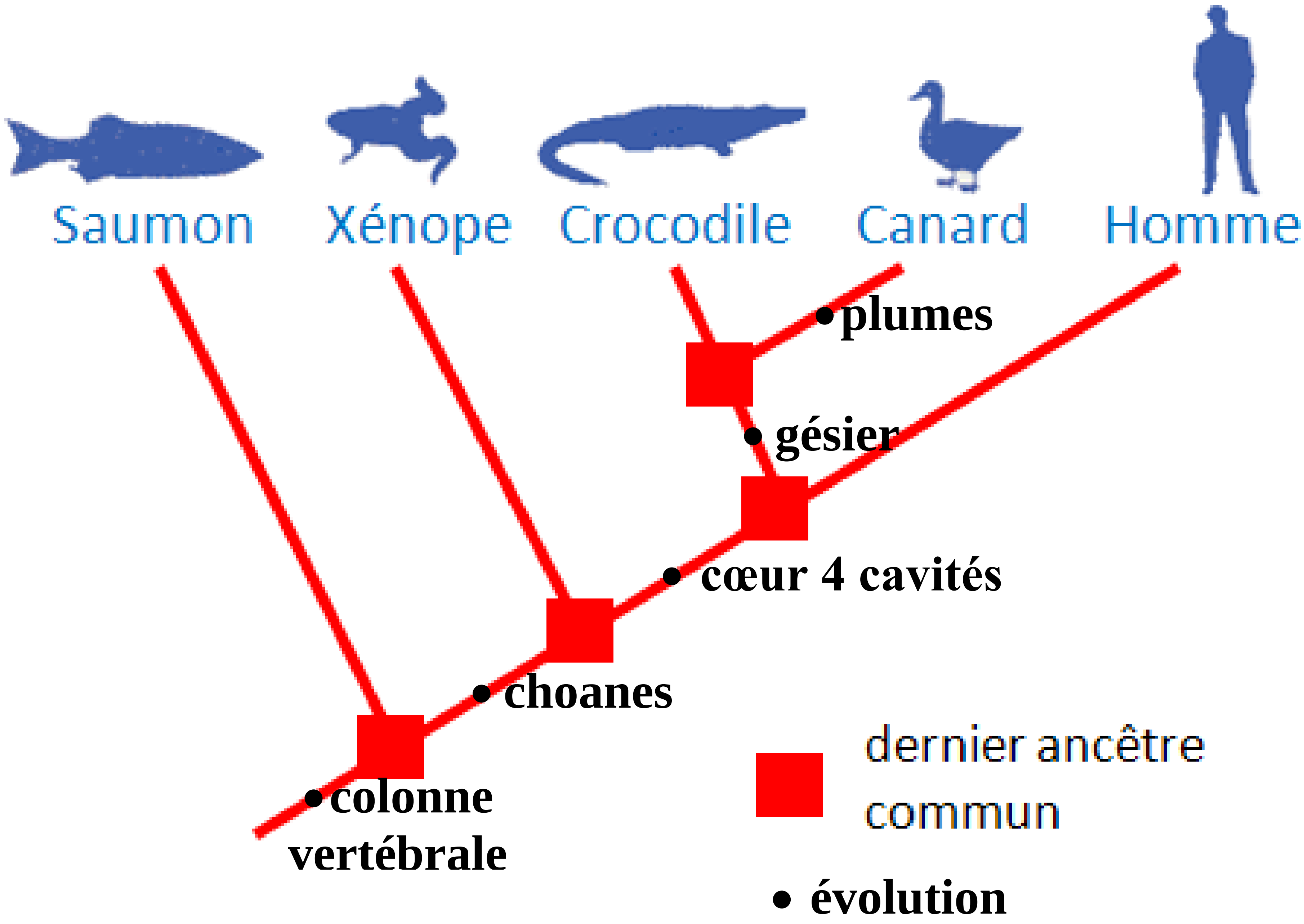
A **= ensemble des arcs (évolutions)**

$d^{+}(y) \geq 2 \ \ (y \in Y)$ **(sauf éventuellement la racine)**

$\forall$ **arc** a **de l'arbre** $\rightarrow$ **valeur** $l(a) \geq 0$

Arbre phylogénétique

- Relations de parenté entre organismes vivants
- Indique qui est proche de qui
- Fondé sur l'analyse de caractères des espèces étudiées
- Branches : transformations de caractères
- Longueur des branches : mesure du changement évolutif (temps écoulé, caractères morphologiques ou moléculaires)



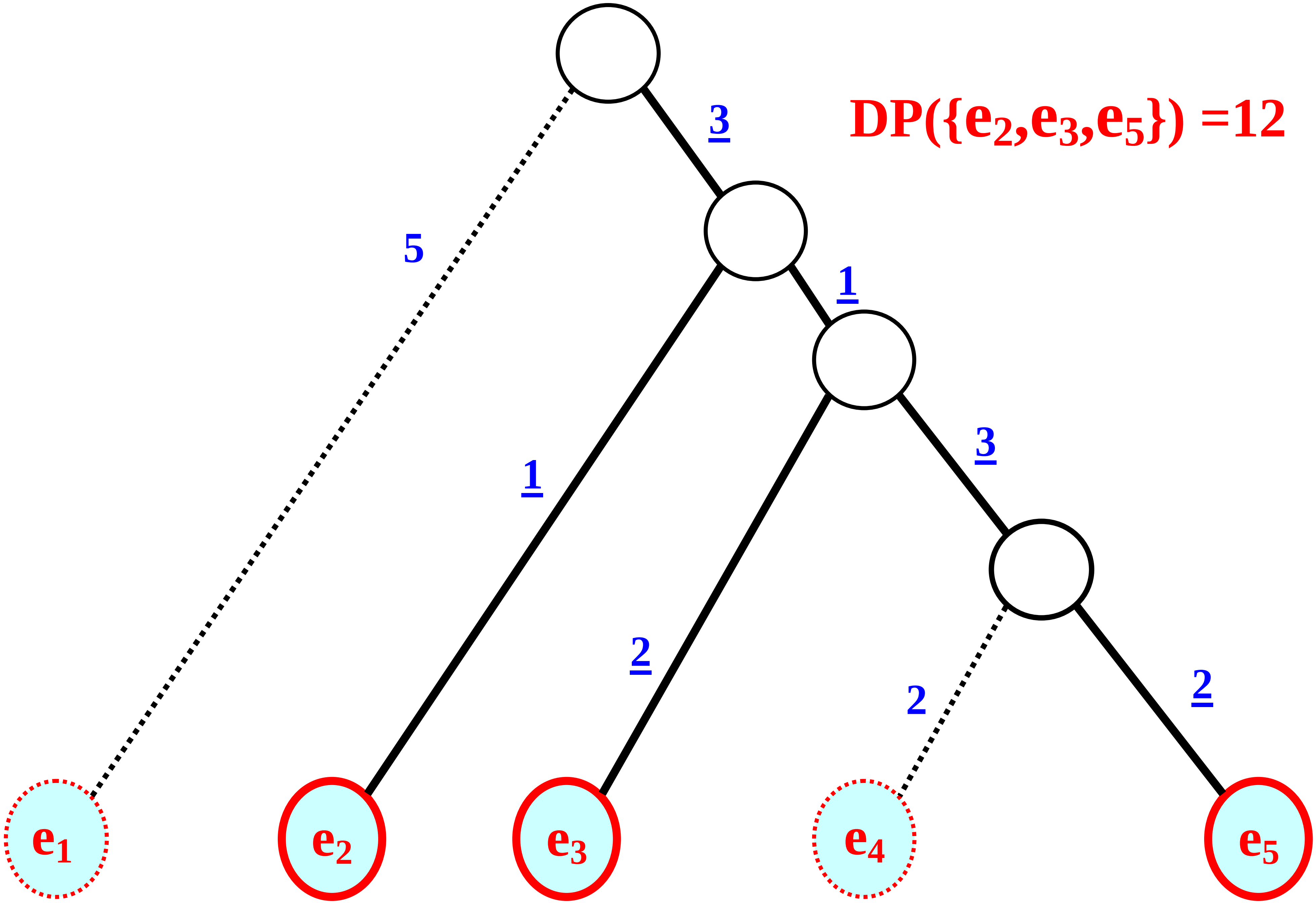
Bergeron et al., Sciences de la vie et de la terre, 2002, Hatier, Paris.

Diversité phylogénétique d'un ensemble d'espèces *S*

DP(*S*)
= somme des valeurs des arcs du sous-arbre minimal connectant les sommets de *S* et la racine

- **Outil rigoureux pour mesurer la biodiversité**
- **Actuellement considérée comme une des principales mesures de la biodiversité d'une collection d'espèces**

Zoological Society of London
World Wide Fund for Nature (WWF)
…



$DP(\{e_2, e_3, e_5\}) = 12$

Un premier problème

Trouver un sous-ensemble de feuilles S (espèces)

- de cardinal K

- qui maximise DP(S)

Algorithme gourmand

- **S = {Ø}**
- **tant que l'ensemble en construction est de cardinal $\leq K$, ajouter le sommet s qui maximise DP$\left(S \cup \{s\}\right)$**

Problème de l' « Arche de Noë »

- **Un arbre phylogénétique**
- **2 décisions possibles pour chaque espèce e_k :**

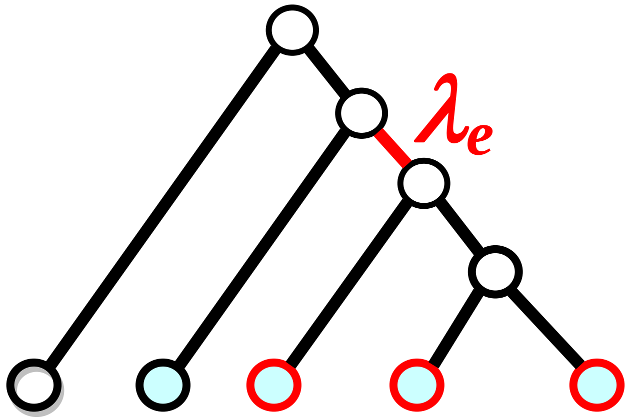
1- Pas d'actions de conservation

$\Rightarrow$ Probabilité de survie = α_k, coût = 0

2 - Actions de conservation

$\Rightarrow$ Probabilité de survie = $\beta_k > \alpha_k$, coût = c_k

Politique à appliquer à chaque espèce pour maximiser l'espérance math. de la diversité phylogénétique ?
(probabilités indépendantes)



probas de survie de e_k

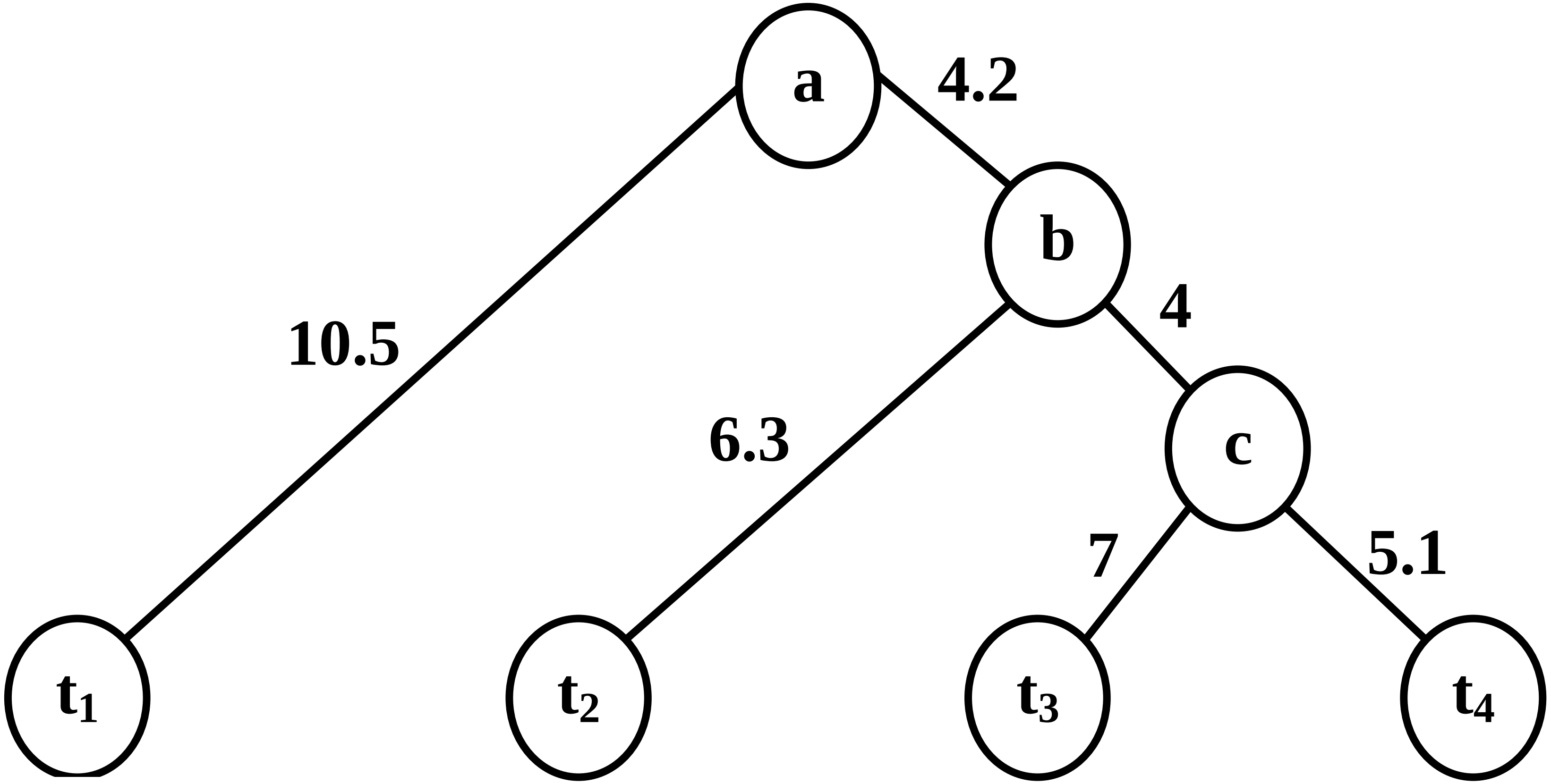
$$\mathrm{Esp(DP)} = \sum_{a \in A} \lambda_a \left(1 - \prod_{k \in F_a \cap S_1} (1 - \beta_k) \prod_{k \in F_a \cap S_2} (1 - \alpha_k) \right)$$

ensemble des arcs de l'arbre

ensemble des feuilles sous l'arc *a*

S_1 : espèces protégées

S_2 : espèces non protégées



probas		probas		probas		probas	
0.5	0.8	0.3	0.6	0.2	0.7	0.3	0.9
coûts		coûts		coûts		coûts	
0	4	0	3	0	7	0	8

$$(10.5 \times 0.5) + (4.2 \times (1 - 0.4 \times 0.8 \times 0.1)) + (6.3 \times 0.6) + \ldots$$

Formulation par un programme mathématique

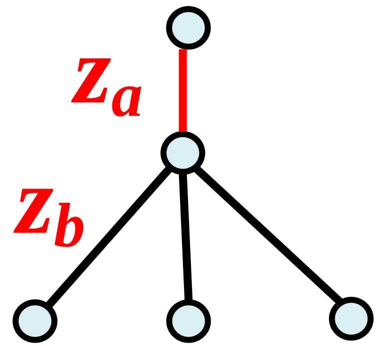
x_k

= 1 ssi actions de conservation appliquées à espèce e_k

$$\begin{cases} \max \displaystyle\sum_{a \in A} \lambda_a \left(1 - \prod_{k \in F_a} \boxed{\left[1 - \alpha_k - x_k (\beta_k - \alpha_k) \right]} \right) \\ \text{s.c.} \left| \begin{array}{ll} \displaystyle\sum_{e_k \text{ espèce}} c_k x_k \leq B & \\ x_k \in \{0,1\} & \forall \text{espèce } e_k \end{array} \right. \end{cases}$$

Probabilité extinction e_k

Reformulation



$$
\begin{cases}
\max \sum_{a \in A} \lambda_a (1 - z_a) \\
z_a = \prod_{b \in A_a} z_b \qquad (a \text{ arc non pendant}) \\
z_a = \boxed{1 - \alpha_k - x_k(\beta_k - \alpha_k)} \qquad (a \text{ arc pendant}) \\
\text{contrainte budgétaire} \\
x_k \in \{0,1\} \qquad \forall \text{espèce } e_k
\end{cases}
$$

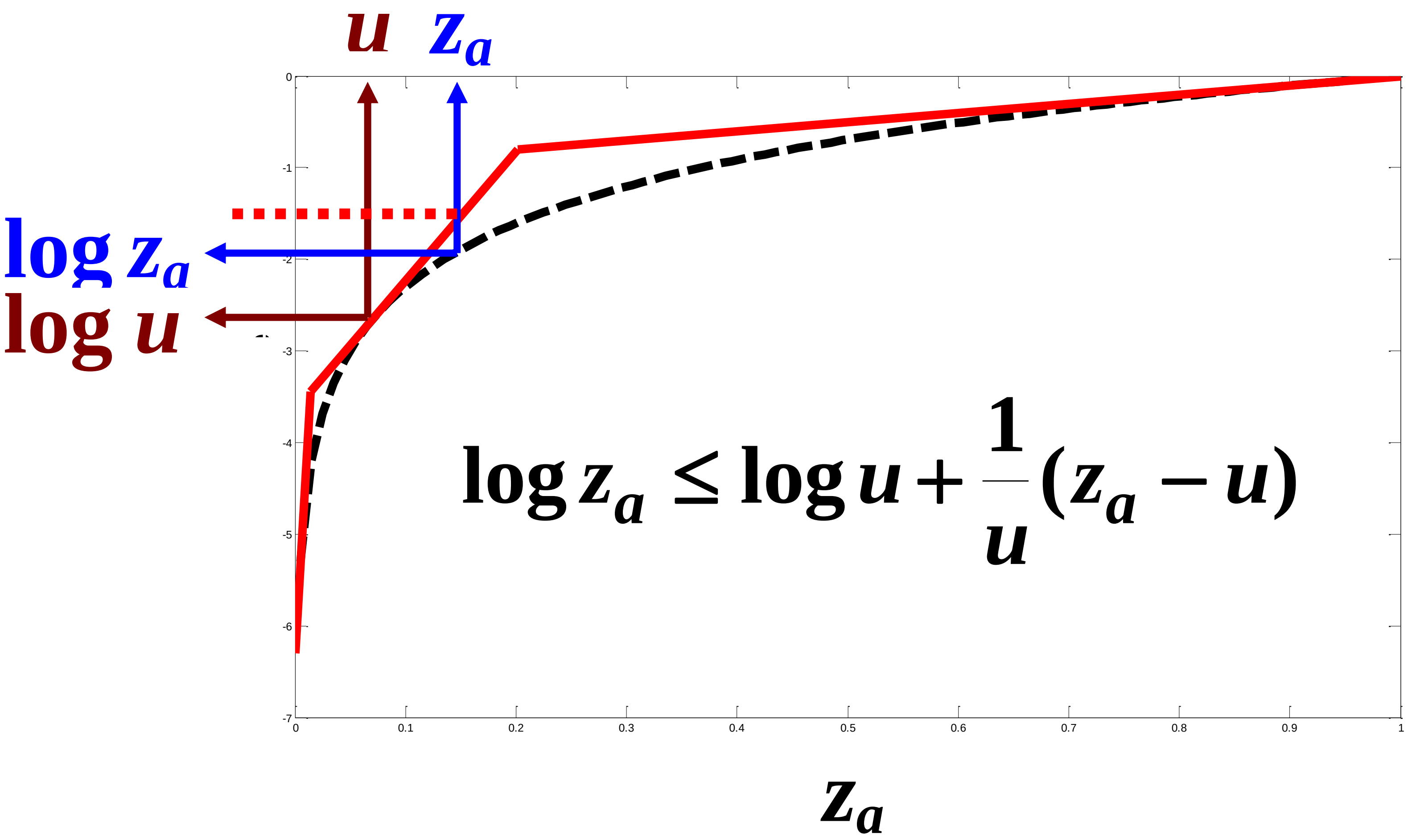
probabilité perte évolution associée à *a*

probabilité extinction e_k

Linéarisation

$$\begin{cases} \max \sum_{a \in A} \lambda_a (1 - z_a) \\ \boxed{y_a = \log z_a} \quad (\forall \text{arc } a) \\ y_a = \sum_{b \in A_a} y_b \quad (\forall \text{arc } a \text{ non pendant}) \\ y_a = x_k \log(1 - \beta_k) + (1 - x_k) \log(1 - \alpha_k) \quad (\forall a \text{ pendant}) \\ \text{contrainte budgétaire} \\ x_k \in \{0,1\} \end{cases}$$

non linéaire


u
z_a
$\log z_a$
$\log u$
0
-1
-2
-3
-4
-5
-6
-7
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
$\log z_a \leq \log u + \frac{1}{u}(z_a - u)$
z_a

$$y_a = \log z_a \Rightarrow y_a \le \log z_a$$

Relaxation :

$$y_a \le \log z_a \longrightarrow y_a \le \log u_k + \frac{1}{u_k}(z_a - u_k) \quad (k \in K)$$

- **Borne supérieure : valeur optimale de la relaxation**

- **Espérance de la diversité phylogénétique :**

$$\sum_{a \in A} \lambda_a \left(1 - \prod_{k \in F_a} (1 - \alpha_k - x_k^*(\beta_k - \alpha_k)) \right)$$

Expérimentation

- **Arbres phylogénétiques avec des milliers d'arcs**
- **de 20 à 50 points u_k pour chaque arête de l'arbre**

  ⇒ **Solutions approchées avec garantie (à moins de 0.1% de l'optimum)**

- **Généralisation : plusieurs stratégies de protection possibles pour chaque espèce**
- **Possibilités de considérer des probabilités de survie non indépendantes : zones protégées**

Conclusion

- **Actuellement érosion très importante de la biodiversité**
- **Optimisation combinatoire : un outil naturel pour aider à prendre des décisions pour enrayer cette érosion**
- **Beaucoup de publications sur le sujet (littérature RO et Conservation Biologique)**
- **Simulations, heuristiques**

- **Certains problèmes bien résolus, d'autres pas**
  ➡ **recherches nécessaires**

- **Gap important entre recherche et mise en œuvre**
  ➡ **collaborations praticiens / chercheurs**

- **Dimension temporelle à considérer**

- **Incertitudes à prendre en compte**

- **Modèles mathématiques :**
  un aspect de la préservation de la biodiversité