



Recherche de similarité dans du code source



<http://www.plade.org/>

Code à analyser

Arbre de syntaxe abstrait (AST)
et normalisé

Intra-projet

Maintenance
et factorisation
du code.
Suivi temporel
de versions.

SortLibrary

QuadraticSortingAlgorithms (QSA)

GnomeSort

```
void gnomeSort(int * a, int n) {  
    int i = 0 ;  
    while (i < n) {  
        if (i == 0 || a[i-1] <= a[i]) i++;  
        else {  
            int tmp = a[i-1] ;  
            a[i-1] = a[i] ; a[i] = tmp ; i-- ; } } }  
...  
BubbleSort  
...
```

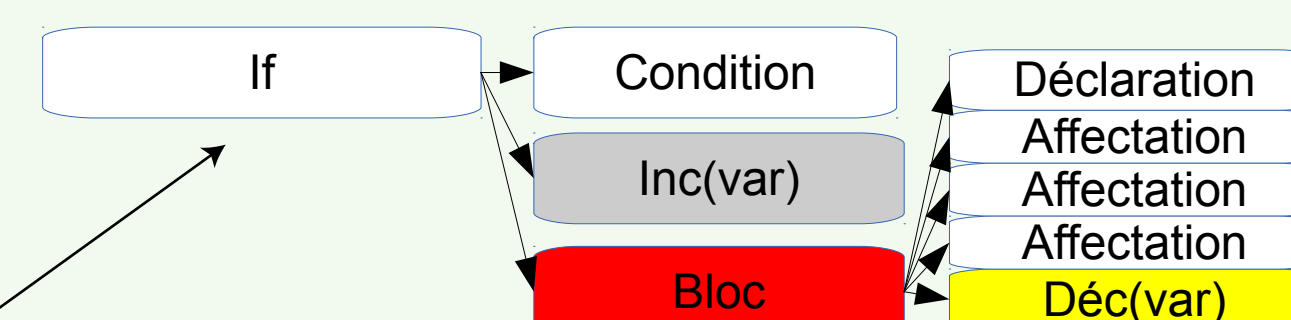


Inter-projets

Copie légitime ou non
du code. Si non légitime :
obfuscation probable.

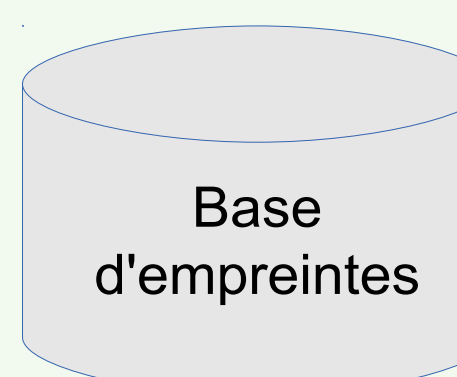
KabouterSort

```
void kabouterSort(int m, int b[]) {  
    unsigned int j;  
    for (j = 1 ; j < n ; ) {  
        if (j > 0 && b[j] < b[j-1]) {  
            swap(b, j, j-1) ; j-- ; }  
        else j = j + 1 ; } }  
  
void swap(int b[], int i, int j) {  
    int tmp = b[i-1/1+0] ;  
    b[i] = b[j] ; b[j] = tmp ; }
```



Empreintes
de sous-arbres
multi-échelle

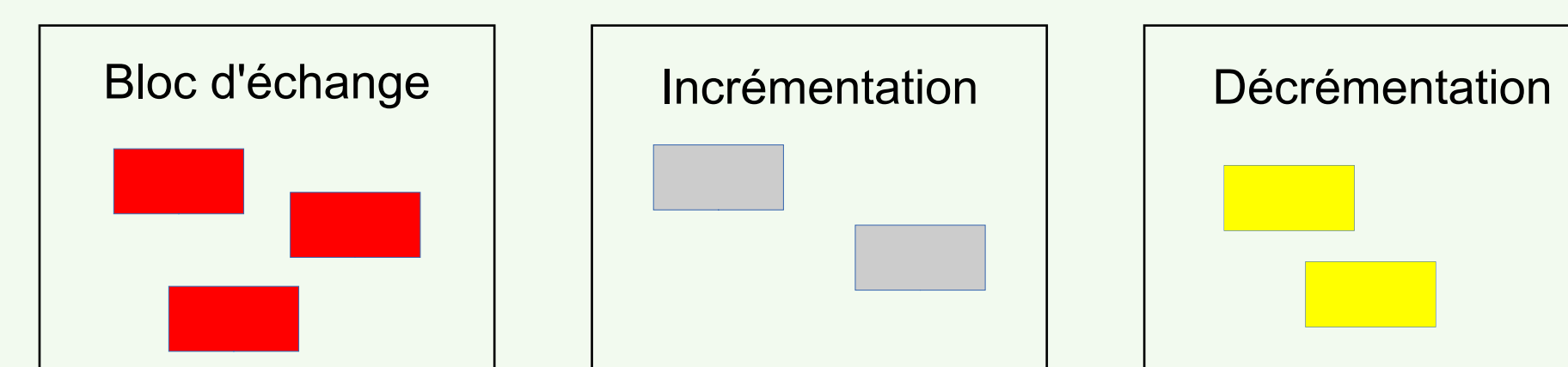
Indexation



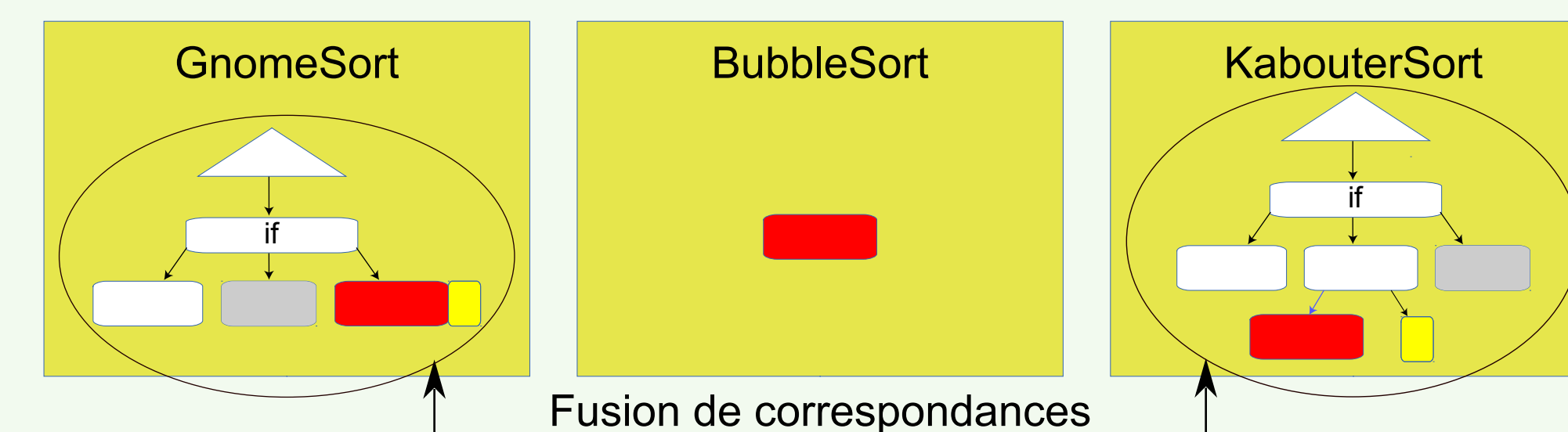
AST normalisé

Bloc d'échange

I - Examen des classes d'équivalence par empreinte



II - Consolidation par proximité spatiale dans l'AST



III - Analyse des correspondances :

- quantification et caractérisation (scores)
- macro-représentation (plans de similarité, tracés dotplot)

	Gnome Sort	Bubble Sort	Kabouter Sort
Gnome Sort		~0,2	~1
Bubble Sort	~0,2		~0,2
Kabouter Sort	~1	~0,2	

Contact : Michel Chilowicz, Etienne Duris, Gilles Roussel