

UNIVERSITÉ PIERRE ET MARIE CURIE

RÉSOLUTION DE PROBLÈMES

Résolution de Mots Croisés par un CSP.

Renaud ADEQUIN
Nadjat BOURDACHE

04/04/2016

1. Modélisation par un CSP et résolution

1. Pour résoudre ce problème, on propose une modélisation qui consiste à associer une variable à chaque mot de la grille. Les mots de la grille étant numérotés dans l'ordre de leur apparition dans la grille (d'abord les mots horizontaux puis les verticaux). On définit ensuite un ensemble de contraintes pour vérifier la cohérence de la grille générée.

Variables :

Pour m mots, on a m variables : Mot_i , $\forall i \in \{1, \dots, m\}$.

Domaine :

Chaque mot de la grille doit appartenir au dictionnaire considéré, notons le $Dict$.

$$D(Mot_i) = \{X \in Dict : |X| = |Mot_i| \}, \forall i \in \{1, \dots, m\} .$$

Contraintes :

- Pour toute paire de mots Mot_i et Mot_j qui se croisent aux positions p pour Mot_i et q pour Mot_j , on définit la contrainte :

$$Mot_i[p] = Mot_j[q].$$

- Pour modéliser le fait qu'un même mot ne peut apparaître plus d'une fois dans la grille, il suffit d'ajouter la contrainte :

$$AllDiff (Mot_1, Mot_2, \dots, Mot_m)$$

2. D'un point de vue algorithmique, nous avons mis en œuvre une classe "Grille". Chaque objet de cette classe représente une instance de mots croisés, et contient une taille, un dictionnaire, une liste de tous les mots de la grille ayant une taille supérieure à 1 ainsi qu'une liste de cases noires.

On pourra à partir d'un objet de cette classe, initialiser une grille à partir d'un fichier texte contenant une grille ou en générer une aléatoirement et la sauvegarder dans un fichier après résolution.

L'interface graphique proposée permet d'ouvrir une grille vide ou partiellement rempli ou d'en générer une en indiquant la taille et le pourcentage de cases noirs. Nous pouvons aussi à tout moment sauvegarder

la grille non/partielement/completement résolu.

3. Nous avons développé les algorithmes AC3 et Forward Checking tel qu'ils ont été définis en cours.

Heuristique :

Nous avons développé plusieurs heuristiques :

- Domaine minimum : qui donne la variable avec le plus petit domaine.
- Max contraintes instance : nous renvoi le mot avec le plus de contraintes avec les variables déjà instanciés.
- Max contraintes : renvoi la variable avec le plus de contraintes.

Afin de déterminer l'heuristique la plus appropriée pour les différents algorithmes, nous avons compté le nombre de mots testés. Après plusieurs tests, nous avons déterminé que la meilleure heuristique pour le Forward Checking était l'heuristique "domaine min".

Cependant il aurait été plus judicieux de combiner les heuristiques en cas d'égalité, mais le sujet nous indiquait de prendre un choix aléatoire pour plus de diversité.

4. En revanche, pour le Conflict BackJumping, afin d'améliorer la résolution, nous avons :
 - Effectué une arc-consistance sur les mots de la grille pour réduire les domaines avant de lancer l'algorithme.
 - Ajouté un Check Forward avant l'appel récursif, ce qui nous permet d'éviter des appels inutiles et d'améliorer encore le temps de résolution des grilles les plus grandes.

Heuristique :

Pour le CBJ, l'heuristique la plus efficace est la même que pour le FC, à savoir "domaine min".

5. Afin d'améliorer les trois algorithmes, nous avons utilisé une structure d'arbre pour sauvegarder et manipuler les dictionnaires et les domaines des différentes variables.

Cette structure permet un parcours rapide du dictionnaire pour l'initialisation des domaines, et surtout, elle permet accélérer les algorithmes de filtrage, puisqu'elle peut permettre d'éliminer plusieurs mots d'un domaine en supprimant des sous arbres de ce dernier, nous évitant ainsi

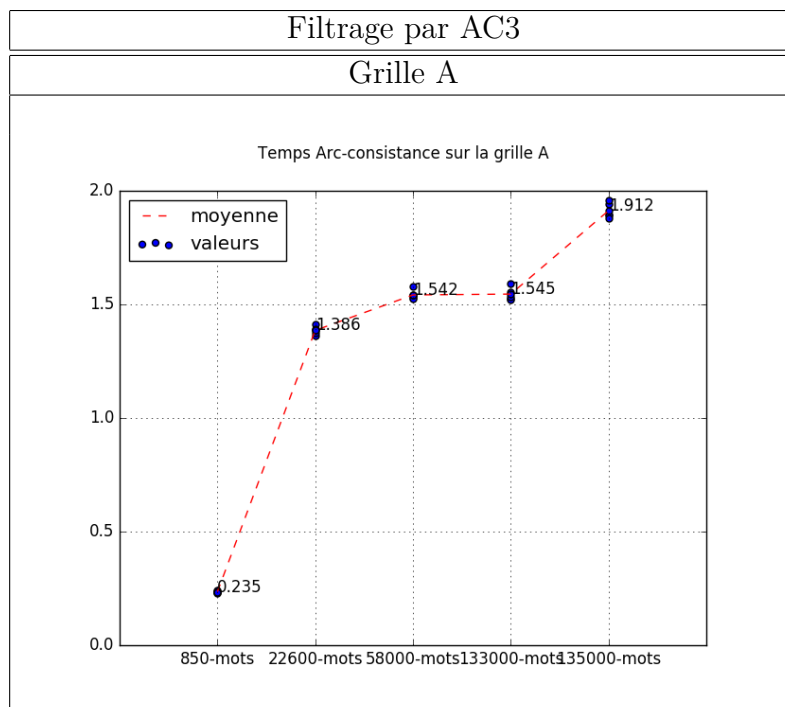
de devoir parcourir et tester chaque mot du domaine.

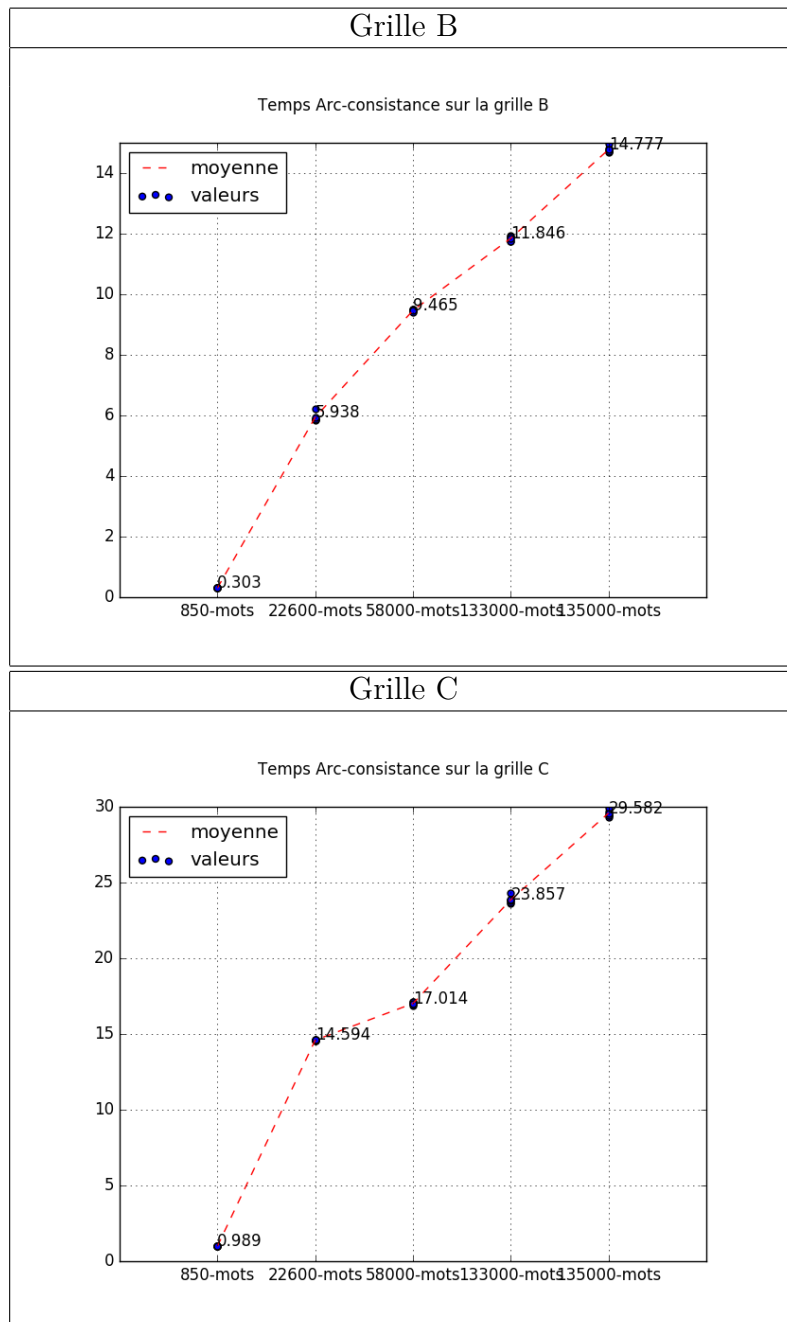
2. Expérimentations

1. Pour le filtrage par AC3, les résultats obtenus pour les différentes grilles sont représentés sur les graphiques ci-dessous.

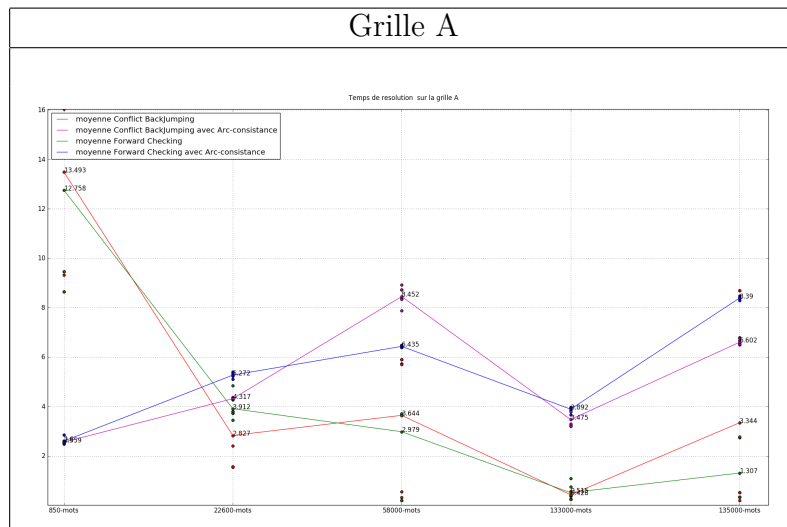
On voit clairement que le temps du filtrage augmente exponentiellement avec la taille des dictionnaires utilisés. Ce qui se justifie par une plus grande taille du domaines des variables, et donc par plus de valeurs pour lesquels on vérifiera la consistance.

On remarque surtout, que plus la grille contient de mots, plus les temps croissent rapidement lorsqu'on augmente la taille des dictionnaires. Ceci est du au fait qu'il y ai plus de variables, et donc plus de contraintes à vérifier et plus de domaines qui croissent lorsque la taille augmente.





2. Pour les algorithmes de résolution FC et CBJ, nous avons comparé les résultats des deux algorithmes avec et sans filtrage et avec différentes tailles de grilles et de dictionnaires.
- Les résultats obtenus sont représenté sur les graphiques suivants.



Pour la grille A, on remarque que le filtrage par AC3 préalable n'est bénéfique que pour les petits dictionnaires. Pour les plus grands dictionnaires, la résolution sans filtrage par AC3 demande un temps de calcul beaucoup plus court que lorsqu'on en effectue un. C'est dû au grand nombre mots à tester dans les domaines des différentes variables.

