

Lecture Automatique de Formules Chimiques

Manuel d'utilisation v1.0

Jean-Yves Ramel - Guillaume Boissier

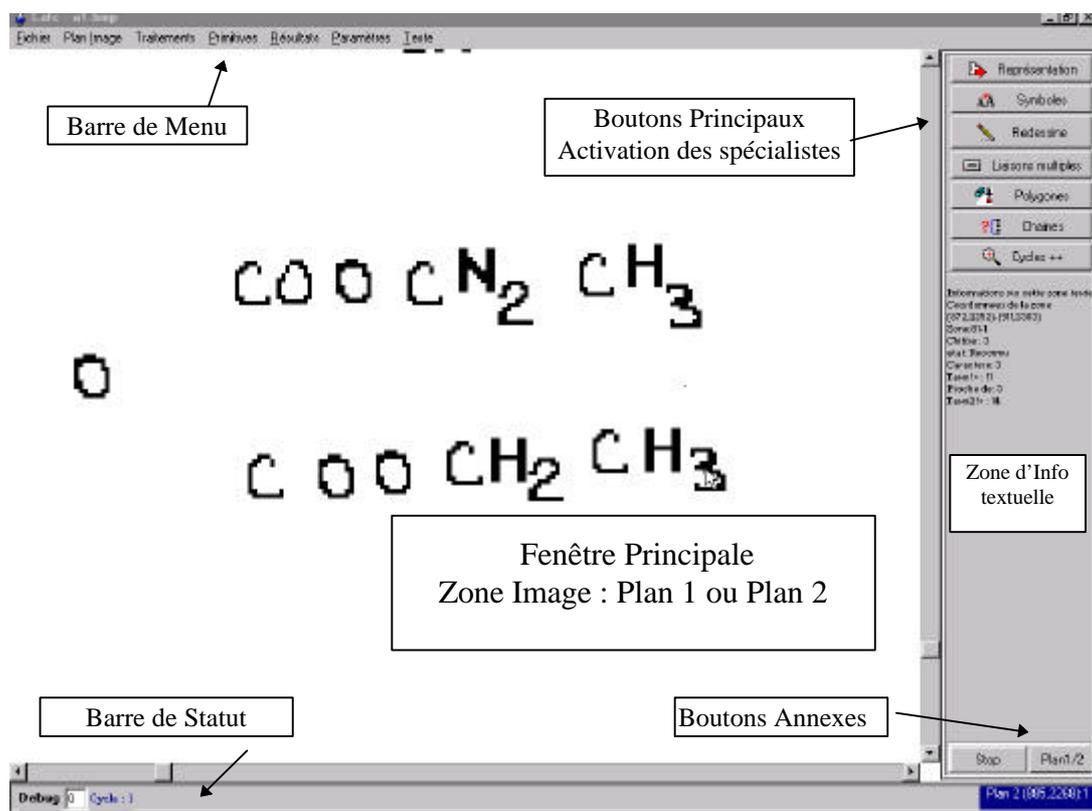
juillet 1998

Avant-propos

La maquette développée, intitulée *Lafc*, fonctionne sous Windows 95 ou NT. Elle a été développée à l'aide du logiciel Borland C++ Builders.

Ce rapport décrit les principales possibilités offertes par ce « logiciel » qui reste pour l'instant une maquette. Cette maquette est une adaptation du logiciel *LAD* (*Lecture automatique de dessins*) réalisé (sous Unix) durant la thèse par JY Ramel.

1. Fenêtre Principale de l'application



La fenêtre principale se décompose en 6 parties :

- Une barre des Menus classique
- Une Zone Image : l'affichage de 2 plans est possible : un plan image initiale (plan 1) et un plan de travail (plan 2).
- Le basculement d'un plan à l'autre se fait grâce au bouton **Plan1/2** de la zone Boutons annexes, le second bouton de cette zone **Stop/Go** sert à arrêter ou à relancer l'analyse en mode debug.
- Une zone Info-Texte pour l'affichage d'informations sous forme textuelle.
- Une Zone Activation des Spécialistes qui est en fait le tableau de bord de l'application.

- Une barre de Statut sur laquelle s'affiche les informations sur l'état courant : Mode-Debug : Oui (1)/Non(0), Plan courant : 1 / 2, barre de progression, ...

A. LA ZONE ACTIVATION DES SPECIALISTES

Il s'agit de boutons qui permettent d'activer les différents spécialistes chargés d'analyser la formule chimique dessinée. Pour avoir plus d'information, voir le rapport technique.

- **Représentation** : Construction de la représentation du contenu de l'image sous forme de Vecteurs et Quadrilatères.
- **Symboles** : Localisation et Reconnaissance du Texte manuscrit.
- **Redessine** : Construction du graphe structurel représentatif des traits constituant la formule chimique et des relations entre ces traits.
- **Liaisons** : détection des liaisons parallèles et croix constituant la formule
- **Polygones** : détection des polygones constituant la formule
- **Chaînes** : Détection des chaînes constituant la formule
- **Totale** : Lancement de l'interprétation complète : évite d'appuyer successivement sur les boutons décrit ci-dessus. Le résultat est affiché dans le plan 2. Ce bouton n'est utilisable que si l'analyse n'a pas encore commencé ; à utiliser juste après l'ouverture de l'image à analyser.

B. LES MENUS

Nous avons développé les menus suivants:

- Menu Fichiers: Ouvrir. Il permet de sélectionner et d'ouvrir le fichier BMP contenant la formule chimique à analyser et reconnaître. L'image est affichée dans le **plan 1** de la fenêtre principale. Le **plan 2** correspond à une «image vierge» servant de zone de travail et d'affichage des résultats fournis par les différents algorithmes.
- Menu Fichiers: Fermer. Fermeture du fichier BMP ouvert et en cours d'analyse. Remise à zéro de l'ensemble des données extraites et variables utilisées afin d'être paré à interpréter une autre image.
- Menu Fichiers: EnregistrerP2. Sauvegarde du contenu du plan 2 (résultat) sous forme d'une image BMP.
- Menu Fichiers: Informations: Affichage d'informations sur l'image en cours d'analyse, notamment le nombre de pixels en x et y.
- Menu Fichiers: Quitter: Fermeture de l'application *Lafc*.
- Menu Plan-Image: ZoomIn / ZoomOut: Zoom avant et zoom arrière sur l'image (plan 1) ou les résultats (plan 2). La taille originale est la taille de zoom maximale *e.i.* zoom avant n'est possible qu'après l'utilisation de zoom arrière. Cette fonction n'est pas optimale et les calculs complexes effectués à chaque nouvel affichage → à utiliser avec modération !
- Menu Plan-Image: Raz plan2: Remise à zéro / effacement des données affichées dans plan 2
- Menu Plan-Image: P2→P1: Recopie du contenu du plan 2 dans le plan 1 → La zone résultat devient la zone à étudier.
- Menu Pré-traitement: Axe médian: Fonction non validée (bug!). Autre méthode utilisable pour obtenir une représentation de l'image sous forme de Vecteurs. Cette méthode n'a pas été retenue.
- Menu Pré-traitement:Restauration: Pour les images de mauvaise qualité (contours en escaliers, trous, points noirs, ...). Amélioration de l'image initiale avant l'analyse. Généralement, cette étape est inutile. Elle doit être combinée avec l'utilisation du menu *Plan-Image :P2→P1* car l'image restaurée se trouve dans le Plan 2 et doit être passée dans le plan 1 pour être analysée.
- Menu Primitives. Ce menu permet d'afficher les différents résultats au cours de l'analyse de l'image. Il est nécessaire d'avoir construit la représentation (bouton **Représentation**) de l'image avant d'essayer d'afficher un résultat. Pour comprendre l'utilité des sous-menus, il est préférable d'avoir lu le rapport technique fourni !!!

Les sous-menu *Vecteurs / Quadrilatères (Squelette) / Composantes connexes* affichent les primitives correspondantes (bruits et formes pleines / Traits / ...) dans le plan 2.

Le sous-menu *Composantes-Texte* affiche dans le plan 2 les composantes Texte localisées après utilisation du bouton **Symboles** de la fenêtre principale.

Le sous-menu *Graphe* affiche dans le plan 2 une représentation du graphe représentatif de l'image construit lors de l'appui sur le bouton **Redessine**.

Le sous-menu *GrapheSélectif* affiche dans le plan 2 une représentation du graphe représentatif de l'image construit lors de l'appui sur le bouton **Redessine**. Cependant, il est ici possible de sélectionner le type des arcs à dessiner parmi ceux existant (L, X, P, S, ...). Voir rapport technique.

Le sous-menu *Quadr.Sélectifs* affiche dans le plan 2 une représentation des Traits contenus dans l'image initiale. Cependant, il est ici possible de sélectionner un type particulier de traits. Seuls les Quadrilatères correspondant à ce type de Traits seront affichés dans le plan 2.

Type associé à un Quadrilatère	Code correspondant
Quadrilatère appartenant à un Polygone	100 + nb de cotés
Quadrilatère appartenant à une Chaîne	200
Quadrilatère appartenant à une Liaison	300
Quadrilatère appartenant à un Caractère	400
Quadrilatère appartenant à une Croix	500 + nb de traits

Types de quadrilatères

- Menu Résultats: Zones Texte. Cette procédure permet d'afficher les zones de texte localisées par le spécialiste dédié à cette tâche. Celui-ci doit donc avoir été activé grace au bouton **Symboles**. Voir aussi le menu *Primitives :Comp.Textes*
- Menu Résultats:Courbes : Inutilisé.
- Menu Résultats:Symboles : Affiche les symboles (liaisons, polygones, chaînes, croix, ...) reconnus par les différents spécialistes chargés de ces tâches :
 - Liaisons → L + nombre de parallèles
 - Polygones → P + nombre de cotés
 - Chaines → C + liste du type 1P11P → 1 liaison simple, suivie d'une liaison multiple, suivie de 2 liaisons simples, suivie d'une liaison multiple.
 - Croix → X + nombre de traits constituant la croix
- Menu Résultats: Texte Reconnu. Cette procédure permet d'afficher les résultats de la reconnaissance de caractères dans le plan n°2.
- Menu Résultats: Sauvegarder Symboles. Ce menu permet l'enregistrement des résultats (symboles chimiques et texte) dans un fichier *.laf. Le format de sauvegarde des données est le suivant :

```

- Fichier Resultats : D:\users\JY\JMBS\dem2.laf -
- (Image Dem2.bmp) -
-=====

#Texte#
Nom=C
NbSommets=2
Sommets=(177,132) (219,170)
Commentaire=

#Texte#
Nom=CH2
NbSommets=2
Sommets=(316,270) (433,324)
Commentaire=

#Symbole#
Nom=Liaison
NbSommets=4
Sommets=(197,239) (197,186) (179,227) (182,183)
Commentaire=L2

```

```

#Symbole#
Nom=Liaison
NbSommets=2
Sommets=(290,287) (225,291)
Commentaire=L1

#Symbole#
Nom=Polygone
NbSommets=6
Sommets=(127,68) (130,127) (96,158) (37,132) (40,83) (81,21)
Commentaire=

#Symbole#
Nom=Chaine
NbSommets=1
Sommets=(171,144)
Commentaire=C1

....

```

- Menu Résultats: Sauvegarder Vectorisation. Ce menu permet l'enregistrement des résultats de la vectorisation (vecteurs, quadrilatères, composantes connexes) dans un fichier texte (*.laf). Rem : les quadrilatères sont transformés en traits par la fonction *Redessine* :
- Menu Texte: Reconnaître. Cette procédure permet la reconnaissance automatique (sans confirmation par l'utilisateur) des zones caractères. Cette procédure nécessite la détection préalable des zones texte. Pour cela cliquer sur « Représentations », puis « Zones Texte » dans la barre d'outils située à droite de l'écran.
- Menu Texte: sous menu Anciennes Méthodes:
 - ◆ Reconnaître. Cette procédure permet la reconnaissance des caractères avec l'assistance de l'utilisateur. En effet celui ci est constamment sollicité pour confirmer les décisions de l'ordinateur. Cette procédure doit être utilisée lors de l'apprentissage du premier document. Cette procédure nécessite la détection préalable des zones texte. Pour cela cliquer sur « Représentations », puis « Zones Texte » dans la barre d'outils située à droite de l'écran.
 - ◆ Ajouter aux modèles. Cette procédure permet l'ajout des zones texte sans toutefois vérifier si les modèles ajoutés étaient réellement nécessaires pour une meilleure reconnaissance. Cette procédure a été très utile lors du développement mais elle ne doit plus être utilisée. Cette procédure nécessite la détection préalable des zones texte. Pour cela cliquer sur « Représentations », puis « Zones Texte » dans la barre d'outils située à droite de l'écran.
 - ◆ Recréer les modèles. Cette procédure permet d'établir une nouvelle base de modèles à partir des caractères détectés sans toutefois vérifier si les modèles ajoutés étaient réellement nécessaires pour une meilleure reconnaissance. Cette procédure a été très utile lors du développement mais elle ne doit plus être utilisée. Cette procédure nécessite la détection préalable des zones texte. Pour cela cliquer sur « Représentations », puis « Zones Texte » dans la barre d'outils située à droite de l'écran.

- Menu Texte: sous menu Etape par Etape :

Ce sous menu contient des procédures permettant de mieux comprendre le fonctionnement de notre algorithme. Leurs exécutions successives aboutissent à une reconnaissance totale, néanmoins elles peuvent être exécutées dans n'importe quel ordre. Pour une meilleure compréhension, il est utile de consulter au préalable le rapport sur la reconnaissance de caractères. Toutes les procédures de ce sous-menu nécessitent la détection préalable des zones texte. Pour cela cliquer sur « Représentations », puis « Zones Texte » dans la barre d'outils située à droite de l'écran.

- ◆ Reconnaissance pure. Cette procédure permet d'observer dans le plan n°2 les résultats obtenus par l'algorithme principal de reconnaissance, utilisant la normalisation et les masques contenus dans la base de modèles.
- ◆ Découpe Zone Doubles. Cette procédure permet d'observer dans le plan n°2 les résultats obtenus par la procédure précédente plus ceux obtenus grâce à l'algorithme de découpage des zones non reconnues.
- ◆ Réduction du doute. Cette procédure permet d'observer dans le plan n°2 les résultats obtenus par les procédures précédentes plus ceux obtenus grâce à l'aide des méthodes secondaires (définies dans notre rapport), à savoir l'étude des quadrilatères significatifs, ainsi que l'étude des demis profils droits.
- ◆ Regroupement par Zone. Cette procédure permet d'observer dans le plan n°2 les résultats obtenus par les procédures précédentes plus ceux obtenus grâce à l'algorithme de localisation des ensembles de zones caractères proches les unes des autres.
- ◆ Détection des chiffres. Cette procédure permet d'observer dans le plan n°2 les résultats obtenus par les procédures précédentes plus ceux obtenus grâce à la détection des zones plus susceptibles de contenir un chiffre de par leur position géographique relative aux zones texte qui les entourent.
- ◆ Analyse. Cette procédure permet d'observer dans le plan n°2 les résultats obtenus par les procédures précédentes plus ceux obtenus grâce à l'analyse sémantique des zones textes. Cette procédure est équivalente à menu Texte, Reconnaître.

- Menu Texte: sous menu Modèles

- ◆ Bibliothèque. Cette procédure permet d'appeler la fenêtre de gestion des masques, et ainsi d'éditer la liste des modèles de la base (voir 2.).

- ◆ Fusionner. Cette procédure permet de demander la fusion des modèles les plus proches contenus dans la base. Cette méthode étant appelée automatiquement dans de nombreux cas il n'est plus utile de l'utiliser depuis ce menu.
- ◆ Changer utilisation. Cette procédure permet de changer le taux d'utilisation de chaque modèle ainsi que le nombre de reconnaissances générales de la base. A utiliser avec précaution.
- ◆ Oublier. Cette procédure permet de supprimer de la base les modèles dont le taux d'utilisation est égal à zéro, et de remettre à zéro tous les taux d'utilisation ainsi que le nombre de reconnaissances générales.
- Menu Texte: sous menu Symboles
 - ◆ Editer la liste. Cette procédure permet d'appeler la fenêtre de gestion des symboles chimiques contenus dans nos dictionnaires et utiles lors de l'analyse sémantique des zones texte (voir 3.).
- Menu Paramètres : Encore inutilisé

2. Fenêtre de gestion de modèles (ou gestion des masques)

Elle est obtenue en cliquant sur Menu Texte: Modèles: Bibliothèque...

- Menu Fichier: Nouveau. Cette procédure permet de créer une nouvelle base à partir de l'éditeur de modèles. Attention lors de la sauvegarde l'ancien fichier « bimdl.dat » sera écrasé.
- Menu Fichier: Ouvrir. Cette procédure permet d'ouvrir la base de modèles contenue dans le fichier « bimdl.dat ».
- Menu Fichier: Editer Police. Cette procédure permet d'éditer la police de caractère utile à l'affichage au sein de la fenêtre principale. Le fichier ouvert est intitulé « police.dat ».
- Menu Fichier: Fermer. Cette procédure permet de fermer et de sauver le fichier en cours d'édition (soit « bimdl.dat » soit « police.dat »).
- Menu Fichier: Quitter. Cette procédure permet de quitter la fenêtre d'édition. Attention les données modifiées ne seront pas enregistrées.
- Menu Edition: Ajouter. Cette procédure permet l'ajout d'un nouveau modèle. Le modèle en cours d'édition devient alors le nouveau modèle.

- Menu Edition: Supprimer. Cette procédure permet de supprimer de la base le modèle courant.
- Menu Edition: Suivant. Cette procédure permet d'éditer le modèle suivant.
- Menu Edition: Précédent. Cette procédure permet d'éditer le modèle précédent.
- Menu Edition: Valider. Cette procédure permet d'enregistrer en mémoire (et non sur disque) les informations contenues dans « Nom du caractère » et « utilisation » situées dans la barre d'outils à droite de l'écran, ainsi que les éventuelles modifications apportées au dessin du modèle.
- Menu Edition: Effacer. Cette procédure permet d'effacer en mémoire (et non sur disque) les informations contenues dans « Nom du caractère » et « utilisation » situées dans la barre d'outils à droite de l'écran, ainsi que le dessin du modèle.
- Menu Aller à: Premier. Cette procédure permet d'éditer le premier modèle de la base.
- Menu Aller à: Dernier. Cette procédure permet d'éditer le dernier modèle de la base.
- Menu Aller à: n ème. Cette procédure permet d'éditer le modèle dont vous entrez le numéro. Si le numéro entré est inférieur à 1, le premier modèle est alors édité. si le numéro entré est supérieur au nombre de modèles contenus dans la base le dernier modèle est alors édité.
- Boite d'outils commande: <<. Identique à Menu Edition Précédent.
- Boite d'outils commande: >>. Identique à Menu Edition Suivant.
- Boite d'outils commande: Valider. Identique à Menu Edition Valider.
- Boite d'outils commande: Effacer. Identique à Menu Edition Effacer.
- Boite d'outils commande: Nom du caractère. Ce champ contient le nom du caractère codé par le modèle en cours d'édition. Toute modification de ce champ doit être validée par la commande Boite d'outils commande: Valider pour être prise en compte.
- Boite d'outils commande: Utilisation. Ce champ contient le nombre de fois où le caractère en cours d'édition a été utilisé pour reconnaître le contenu d'une zone texte. Toute modification de ce champ doit être validée par la commande Boite d'outils commande: Valider pour être prise en compte.
- Boite d'outils commande: N° du caractère. Ce champ permet de visualiser l'emplacement dans la liste du caractère en cours d'édition. Une modification

(même après validation) n'entraînera aucun changement (et sera totalement ignorée).

- Boîte d'outils Masque. Tous les boutons numérotés (de 0 à 79) correspondent à une case du damier codant le modèle. Ces boutons permettent de modifier le dessin d'un caractère. Toute modification doit être validée par la commande Boîte d'outils commande: Valider pour être prise en compte.

3. Fenêtre de Gestion des symboles

Elle est obtenue en cliquant sur menu Texte: Symboles: Editer la liste...

- Menu Fichier: Ouvrir. Cette commande permet l'édition de la liste des symboles chimiques contenus dans nos dictionnaires.
- Menu Fichier: Fermer. Cette commande permet de sauvegarder et de fermer les différents fichiers contenant les symboles édités. Lors de la sauvegarde les symboles sont triés, permettant ainsi une utilisation simplifiée pour l'utilisateur.
- Menu Fichier: Quitter. Cette commande permet de quitter la fenêtre et de fermer les fichiers en cours d'édition sans enregistrer les modifications.
- Menu Edition: Valider. Cette procédure permet de sauvegarder en mémoire les modifications effectuées sur le symbole en cours d'édition.
- Menu Edition: Ajouter. Cette procédure permet l'ajout d'un nouveau symbole au dictionnaire. Le symbole en cours d'édition devient alors le nouveau symbole.
- Menu Edition: Supprimer. Cette commande permet la suppression du symbole en cours d'édition.
- Menu Aller à: Suivant. Cette procédure permet d'éditer le symbole suivant de nos dictionnaires.
- Menu Aller à: Précédent. Cette procédure permet d'éditer le symbole précédent de nos dictionnaires.
- Menu Aller à: Premier. Cette procédure permet d'éditer le premier symbole.
- Menu Aller à: Dernier. Cette procédure permet d'éditer le dernier symbole.
- Menu Aller à: n ème. Cette procédure permet d'éditer le modèle dont on a saisi le numéro.

- Bouton <<. Effet identique à menu Aller à: Précédent.
- Bouton >>. Effet identique à menu Aller à: Suivant.
- Bouton Sauver & Quitter. Menu Fichier Fermer suivit de menu Fichier Quitter.
- Bouton Valider. Effet identique à menu Edition: Valider.
- Bouton Supprimer. Effet identique à menu Edition: Supprimer.
- Champ Nom. Contient le nom du symbole chimique en cours d'édition. Toute modification doit être validée pour être prise en compte.
- Champ N°. Contient le numéro du symbole en cours d'édition. Toute modification sera ignorée (cet attribut étant géré automatiquement) .