

Un framework organisationnel pour la conception et l'implantation multi-agent de métaheuristiques

D. Meignan J.-C. Créput A. Koukam
david.meignan@utbm.fr jean-charles.creput@utbm.fr abder.koukam@utbm.fr

Laboratoire Systèmes et Transport,
Université de Technologie de Belfort-Montbéliard, France

Résumé

Les travaux autour des métaheuristiques ont récemment évolué vers de nouveaux enjeux tels que : la simplicité, la robustesse et la modularité des métaheuristiques. L'intelligence artificielle distribuée et plus particulièrement les systèmes multi-agent semblent être des approches prometteuses pour y répondre. Ainsi, dans cet article une approche multi-agent pour la modélisation et l'implantation de métaheuristiques est adoptée. Dans ce cadre, nous proposons AMF (Agent Metaheuristic Framework) un framework organisationnel et multi-agent pour la modélisation et l'implantation de métaheuristiques. Ce framework fournit un modèle et des lignes méthodologiques permettant la modélisation de métaheuristiques et facilitant la distribution et l'intégration de composants d'apprentissage au sein des métaheuristiques. Afin d'illustrer l'utilisation d'AMF et de l'apprentissage artificiel dans les métaheuristiques, nous présentons une métaheuristique originale fondée sur le principe de coalition d'agents.

Mots-clés : optimisation combinatoire, métaheuristique, système multi-agent

Abstract

Research in metaheuristics has recently evolved to new issues such as : simplicity, robustness and modularity of metaheuristics. Distributed Artificial Intelligence and particularly multi-agent systems seem to be a promising field of research to tackle these new issues. In this paper we propose AMF, an Agent Metaheuristic Framework that aims at supporting the design and implementation of metaheuristics using an organizational and agent-oriented approach. This framework gives a model and a set of methodological guidelines to support the design of metaheuristics and facilitate the distribution and integration of artificial learning components. This paper also introduces an original metaheuristic called Coalition-Based Metaheuristic (CBM) that illustrates the use of AMF and artificial learning in metaheuristics.

Keywords: combinatorial optimization, metaheuristic, multi-agent system

1 Introduction

Les métaheuristiques sont des méthodes permettant la résolution approchée de problèmes d'optimisation combinatoire. Ces méthodes présentent un intérêt certain pour la résolution de problèmes d'optimisation de grande taille où une solution doit être trouvée en un temps restreint. Les travaux autour des métaheuristiques ont récemment évolué vers de nouveaux enjeux. Cela concerne non seulement la recherche de performances face à des problèmes à large échelle ou des problèmes dynamiques, mais aussi la recherche de méthodes simples, robustes et modulaires. Ces nouveaux aspects ont été mis en avant dans différentes études [1, 2] où il ressort que la simplicité, la robustesse et la modularité constituent des critères importants pour une utilisation effective des métaheuristiques. L'intelligence artificielle distribuée et plus particulièrement les systèmes multi-agents (SMA) sont des approches prometteuses pour répondre à ces nouveaux enjeux.

Les SMA sont naturellement bien adaptés à la modélisation de métaheuristiques. Ces deux domaines exploitent en effet la métaphore sociale et le paradigme de l'auto-organisation [3]. De plus, les SMA sont largement utilisés dans le cadre des métaheuristiques à base de population, hybrides ou distribuées. Cependant, l'usage des SMA dans le cadre des métaheuristiques manque encore d'outils de modélisation et d'implantation. Dans cet article nous proposons AMF (Agent Metaheuristic Framework), un framework organisationnel et multi-agent pour la modélisation et l'implantation de métaheuristiques. Ce framework fournit un modèle organisationnel de métaheuristique permettant de décrire une métaheuristique à partir de quatre rôles fondamentaux : diversifieur, intensifieur, guide et stratège. À partir de ce modèle, une métaheuristique particulière peut être vue

comme un raffinement du modèle organisationnel d'AMF. De plus, AMF apporte un ensemble de lignes méthodologiques permettant de modéliser une métaheuristique sous la forme d'un SMA. Le but de cette approche est d'encourager la modularité, la réutilisation et de faciliter la distribution de métaheuristicques.

Afin d'illustrer l'utilisation d'AMF, cet article introduit une métaheuristique nommée "Coalition-Based Metaheuristic" (CBM). Dans CBM plusieurs agents organisés en coalition traitent simultanément un problème d'optimisation. Chaque agent dispose d'un ensemble d'opérateurs permettant d'explorer l'espace des solutions du problème. Ils améliorent individuellement leur stratégie de recherche par apprentissage par renforcement et coopèrent par mimétisme et échange de solutions. L'efficacité de CBM et l'intérêt des mécanismes d'apprentissage est soulignée par une application à un problème de tournées de véhicules (VRP, Vehicle Routing Problem).

Cet article La section 2 présente le modèle organisationnel ainsi que les lignes méthodologiques d'AMF. La section 3 introduit la métaheuristique à base de coalition. Son application au problème de tournées de véhicules ainsi que les résultats expérimentaux sont détaillés dans la section 4.

2 AMF : Un framework pour la modélisation et l'implantation de métaheuristicques

Cette partie introduit AMF (Agent Metaheuristic Framework) un framework pour la modélisation et l'implantation de métaheuristicques en adoptant une approche organisationnelle et multi-agent. Ce framework est fondé sur un modèle organisationnel de métaheuristique. Dans un premier temps, nous présentons le métamodèle RIO (Rôle Interaction Organisation) [4] sur lequel repose le modèle organisationnel de métaheuristique d'AMF. Le modèle organisationnel d'AMF puis les lignes méthodologiques associées sont ensuite détaillés.

2.1 Métamodèle RIO

Le métamodèle RIO [4] s'articule autour de trois concepts clés : le rôle, l'interaction et l'organisation. Un rôle est l'abstraction d'un comportement ou d'un statut au sein d'une organisation. Une interaction lie deux rôles de manière à

ce qu'une action du premier rôle engendre une réaction du second rôle. Une organisation est définie comme un ensemble de rôles en interaction. À partir de ces trois concepts, un agent est défini comme une entité communicante jouant un ou plusieurs rôles. De manière plus générale, un rôle peut être associé à un ou plusieurs agents et un agent peut jouer un ou plusieurs rôles. À partir de ce métamodèle, nous décrivons dans la partie suivante une métaheuristique en terme de rôles, d'interactions et d'organisations.

2.2 Description organisationnelle d'une métaheuristique

D'un point de vue organisationnel, une métaheuristique peut être définie comme une organisation dont l'objectif est d'explorer l'espace de recherche d'une instance de problème afin de trouver une solution optimale ou proche de l'optimale. Cette exploration combine, d'une part, des mécanismes d'intensification, et d'autre part, des mécanismes de diversification. L'intensification permet de concentrer la recherche dans les zones de l'espace de recherche prometteuses. La diversification sert à explorer de nouvelles zones de l'espace de recherche. Ces deux tendances sont guidées par un ensemble d'informations structurées relatif à l'espace de recherche. Ces informations servent à coordonner et à équilibrer l'intensification et la diversification. De plus, la stratégie de recherche peut être adaptée en fonction du contexte d'optimisation et des expériences de recherche.

À partir de cette définition nous pouvons identifier au sein d'une métaheuristique quatre rôles fondamentaux que nous nommerons : Intensifieur, Diversifieur, Guide et Stratège.

Le rôle intensifieur a pour objectif de concentrer la recherche dans les zones prometteuses de l'espace de recherche. L'intensification est une exploration locale de l'espace de recherche permettant de trouver de nouvelles meilleures solutions à partir d'une solution, d'un ensemble de solutions ou d'informations restreignant l'espace de recherche.

Le rôle diversifieur permet d'examiner les zones de l'espace de recherche encore non explorées. La diversification est une exploration, souvent aléatoire et éventuellement fondée sur des informations concernant les zones de l'espace de recherche déjà explorées, qui fournit de nouveaux points de départ pour l'intensification.

Le rôle guide coordonne les rôles *intensifieur* et *diversifieur* et équilibre le rapport entre les tendances d'intensification et de diversification. Ce rôle centralise les informations issues de l'intensification et de la diversification au sein d'une mémoire. Nous utilisons le terme "mémoire" en référence au principe de mémoire adaptative de l'approche AMP [5]. Ainsi, le rôle *guide* est chargé d'appliquer une stratégie d'exploration fondée sur un composant de mémoire en utilisant les deux rôles précédents (*intensifieur* et *diversifieur*).

Le rôle stratège est chargé d'ajuster ou de changer la stratégie de recherche mise en place par les rôles *guide*, *intensifieur* et *diversifieur* en fonction du contexte d'optimisation et éventuellement des expériences. Ce rôle correspond aux mécanismes d'adaptation et d'auto-adaptation dans les métaheuristiques. Ce rôle n'est pas systématiquement instancié dans les métaheuristiques.

La figure 1 représente le modèle organisationnel de métaheuristique. À partir de ce modèle, il est possible de décrire plusieurs métaheuristiques. Celles-ci se distinguent par la manière de réaliser les différents rôles (intensifieur, diversifieur, guide et stratège) et par les informations échangées au travers des différentes interactions.

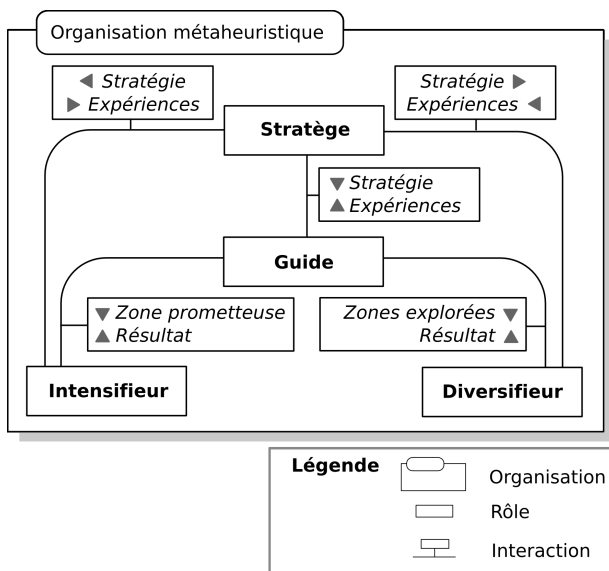


FIG. 1 – Modèle organisationnel de métaheuristique

Le tableau 1 présente la réalisation des quatre rôles du modèle AMF pour six métaheuristiques classiques : recuit simulé, recherche tabou, recherche locale itérée, recherche à voisinage variable, algorithme génétique et algorithme de

colonie de fourmis.

Dans ce tableau, la deuxième et la troisième colonne correspondent aux rôles intensifieur et diversifieur. Pour certaines métaheuristiques, bien que l'on puisse distinguer les tendances d'intensification et de diversification, un seul mécanisme réalise ces deux rôles. C'est le cas pour la recherche tabou, le recuit simulé et les algorithmes de colonie de fourmis. La quatrième colonne correspond aux mécanismes relatifs au rôle guide. La dernière colonne présente une liste non exhaustive de mécanismes d'adaptation. Pour la recherche tabou, l'extension de recherche tabou réactive [6] introduit un mécanisme d'adaptation dynamique de la taille de la liste tabou. Dans le cas du recuit simulé, la méthode de recuit simulé adaptatif [7] permet de modifier automatiquement la loi de refroidissement. Plusieurs extensions adaptatives des algorithmes génétiques existent. Nous avons fait apparaître dans le tableau les trois principaux objectifs de ces extensions décrites dans [8].

2.3 Lignes méthodologiques pour la modélisation de métaheuristiques

Le modèle organisationnel d'AMF permet d'obtenir une métaheuristique sous la forme d'un système multiagent en suivant un processus de conception inspiré de la démarche de conception associée à RIO [4]. Quelques lignes méthodologiques permettant de concevoir un système multiagent associé à une métaheuristique à partir du modèle organisationnel de métaheuristique, sont présentées ci-après.

Le processus de conception se compose de trois étapes : (1) Raffinement du modèle organisationnel d'AMF, (2) Agentification, (3) Spécialisation de la métaheuristique.

La première étape consiste à raffiner le modèle organisationnel d'AMF afin de détailler les comportements des rôles ainsi que les interactions. Le résultat de cette étape est la description d'une organisation de métaheuristique particulière où l'ensemble des rôles et des interactions est décrit. Il est possible dans cette étape de réutiliser des rôles précédemment décrits, notamment pour concevoir une métaheuristique hybride.

La seconde étape a pour objectif de décrire la structure du système multiagent associé à la métaheuristique. Pour cela il est nécessaire d'identifier les différents types d'agents composant le système multiagent, de préciser l'affectation

TAB. 1 – Réalisation des rôles du modèle organisationnel pour différentes métaheuristiques

Métaheuristique	Rôle intensifieur	Rôle diversifieur	Rôle guide	Rôle stratège
Recuit simulé	Recherche dans le voisinage		Critères d'acceptation (température) et refroidissement	Adaptation de la loi de refroidissement [7]
Recherche tabou	Recherche dans le voisinage contraint par la liste tabou		Mise à jour de la liste tabou	Adaptation de la taille de la liste tabou [6]
Recherche locale itérée	Recherche locale	Perturbation	Critères d'acceptation	
Algorithme génétique	Croisement	Mutation	Sélection	Adaptation de la taille de la population, du taux de mutation et de la probabilité de croisement [8].
Algorithme de colonie de fourmis	Création de solutions par les fourmis		Mise à jour des traces de phéromone et évaporation	

des rôles aux agents et de décrire l'ordonnement des rôles pour chaque type d'agent. La principale donnée d'entrée pour effectuer cette étape correspond au modèle raffiné de métaheuristique précédemment obtenu.

La troisième et dernière étape consiste à spécialiser le système multiagent pour le traitement d'un problème d'optimisation particulier. Par exemple, si un algorithme génétique a été décrit sous la forme d'un système multiagent, la spécialisation correspond à la définition des opérateurs de mutation et de croisement spécifiques aux problèmes d'optimisation traités.

3 Métaheuristique à base de coalition

Cette section illustre l'usage d'AMF dans le cadre de la conception d'une métaheuristique nommée CBM (Coalition-Based Metaheuristic). CBM est une méthode fondée sur la métaphore de la coalition. Le terme "coalition" est ici utilisé pour désigner un SMA où les agents disposent de capacités de décision et coopèrent par le biais d'interactions directes [9]. La partie 3.1 donne le principe général de CBM, puis les parties 3.2 et 3.3 suivent les lignes méthodologiques d'AMF pour détailler l'architecture de CBM. Les deux dernières parties 3.4 et 3.5 s'intéressent aux mécanismes de décision et d'apprentissage mis en œuvre.

3.1 Principe

Face à un problème d'optimisation combinatoire que l'on souhaite résoudre à l'aide des

métaheuristiques, il est souvent possible d'utiliser différentes structures de voisinage pour effectuer la recherche. Plus généralement on dispose souvent d'un large panel d'opérateurs permettant d'effectuer une recherche de solutions dans l'espace de solutions. Nous nommerons "opérateur", les procédures permettant d'obtenir une ou plusieurs solutions à partir d'un ensemble éventuellement vide de solutions initiales. Ainsi, un opérateur peut être : une méthode de construction de solution, une procédure de recherche ou de descente locale, ou encore, un opérateur de mutation ou croisement issu des algorithmes génétiques. La bonne utilisation de ces opérateurs est un point crucial pour l'efficacité de la métaheuristique. Ainsi, un algorithme de recherche doit permettre d'appliquer le bon opérateur au bon moment.

Nous proposons au travers de CBM une méthode originale permettant de conditionner le choix des opérateurs au contexte d'optimisation à l'aide d'un composant de décision. Ce dernier est chargé d'analyser le contexte d'optimisation et d'effectuer le choix d'un opérateur à appliquer sur une solution ou un ensemble de solutions. De plus, les règles de décision sont modifiées au cours de l'optimisation par apprentissage.

Dans CBM, la recherche est réalisée par un ensemble d'agents organisés sous forme de coalition. Nous utilisons le terme de coalition pour décrire un système multiagent où les agents sont identiques, ont des capacités de décision et coopèrent par le moyen d'interactions directes. Dans notre cas, la coalition est composée d'agents capables d'effectuer indépendamment une recherche de solution à l'aide d'opé-

rateurs et d'adapter leur stratégie de recherche par apprentissage. La coopération entre agents prend deux formes : tout d'abord, les agents s'informent des zones prometteuses de l'espace de solution, ensuite, les agents s'échangent des informations sur les stratégies de recherche performantes.

Ainsi, les principales caractéristiques de notre approche à base de coalition sont : (i) l'utilisation d'un mécanisme de décision permettant de conditionner le choix d'un opérateur au contexte d'optimisation, (ii) l'apprentissage de règles de décision au cours de l'optimisation, (iii) la coopération entre agents par échange d'informations sur l'espace de recherche ainsi que sur les stratégies. Dans les parties suivantes, les lignes méthodologiques définies dans le framework AMF guideront le processus de modélisation de la métaheuristique CBM.

3.2 Raffinement du modèle organisationnel d'AMF

Pour décrire les principaux composants de CBM le modèle organisationnel d'AMF est raffiné. La figure 2 présente l'organisation CBM résultant du raffinement.

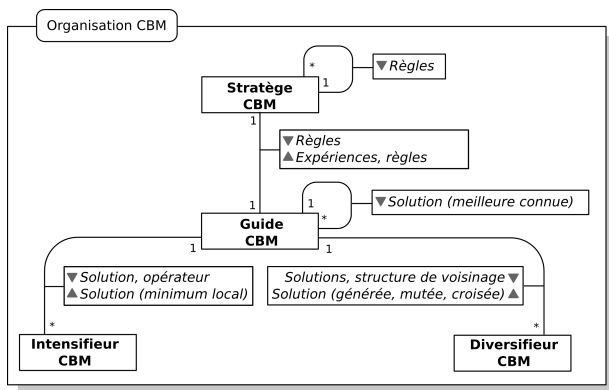


FIG. 2 – Raffinement du modèle organisationnel d'AMF pour CBM

Rôle intensifieur : L'intensification est réalisée par un ensemble d'opérateurs de descente locale. Ainsi, le rôle intensifieur dans l'organisation CBM est chargé d'obtenir un minimum local suivant un voisinage donné à partir d'une solution. Lors de la réception d'une requête, le rôle intensifieur applique une descente locale suivant la structure de voisinage spécifiée et retourne l'optimum local obtenu.

Rôle diversifieur : La diversification est réalisée par des opérateurs de construction, muta-

tion et croisement. On considère ici le croisement comme étant un moyen de diversifier étant donné que contrairement aux opérateurs de descente locale, les opérateurs de croisement ne garantissent pas d'obtenir une meilleure solution. Ainsi, le rôle diversifieur dans l'organisation CBM est chargé de fournir une nouvelle solution à partir d'un ensemble de solutions en appliquant soit une méthode de construction, soit une mutation, soit un croisement.

Rôle guide : Le rôle guide dans l'organisation CBM est chargé, de choisir et de faire appliquer les opérateurs, et d'informer les autres guides de la meilleure solution connue. Pour cela le rôle guide gère une solution courante, la meilleure solution trouvée par l'agent depuis le début de l'optimisation, et la meilleure solution connue au sein de la coalition. La stratégie de recherche mise en place par le rôle guide est issue d'un mécanisme de décision permettant à partir du contexte d'optimisation de sélectionner l'opérateur à appliquer. Le système de décision est détaillé dans la partie 3.4. Le comportement du rôle guide CBM consiste, à observer le contexte d'optimisation, choisir un opérateur puis interagir avec le rôle intensifieur ou diversifieur pour appliquer l'opérateur. Si la solution obtenue est la meilleure connue, alors les autres instances du rôle guide sont informées de cette solution.

Rôle stratège : Le rôle stratège dans l'organisation CBM observe les expériences de recherche issues du rôle guide et modifie en conséquence les règles de décision appliquées par le rôle guide. Ceci est effectué par apprentissage par renforcement et par mimétisme. L'apprentissage par renforcement consiste à favoriser les actions ayant permis d'obtenir de bonnes solutions. L'apprentissage par mimétisme consiste à observer les autres rôles stratèges et de s'inspirer des stratégies bénéfiques pour modifier les règles de décision. Ces deux mécanismes d'apprentissage sont détaillés dans la partie 3.5. Le comportement du rôle stratège consiste à observer le rôle guide. Lorsque la meilleure solution trouvée est améliorée, le rôle stratège applique un apprentissage par renforcement, et lorsque la meilleure solution connue est améliorée, une procédure de mimétisme est appliquée.

3.3 Agentification

À partir du modèle organisationnel de CBM plusieurs architectures de système multi-agent peuvent être envisagées. Dans CBM chaque agent joue l'ensemble des rôles. Ainsi, les

agents de la coalition sont capables de réaliser une optimisation indépendamment des autres agents. L'ajout ou la suppression d'un agent ne remet pas en cause la réalisation de l'optimisation. Cependant, les interactions entre agents permettent d'obtenir de meilleurs résultats. L'agentification de l'organisation CBM est illustrée par la figure 3.

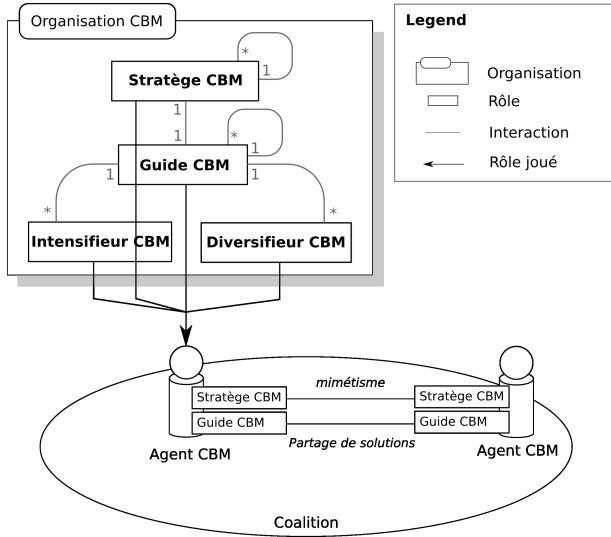


FIG. 3 – Agentification de l'organisation CBM

Le comportement d'un agent dans la coalition consiste à jouer simultanément les quatre rôles. Deux types d'interactions entre agents sont présents au sein de la coalition. La première interaction est réalisée par le rôle guide et consiste à échanger les meilleures solutions connues. La seconde interaction est effectuée par le rôle stratège dans le cadre du mimétisme.

Avant de décrire la spécialisation de CBM pour traiter un problème de tournées de véhicules, le mécanisme de décision ainsi que les mécanismes d'apprentissage qu'elle intègre sont détaillés dans les deux parties suivantes.

3.4 Système de décision de CBM

L'objectif du système de décision dans CBM est de sélectionner un opérateur à appliquer en fonction du contexte d'optimisation. Ce système est à la base du rôle guide permettant de coordonner l'intensification et la diversification. Le système de décision utilisé est issu de l'approche "Adaptive Large Neighborhood Search" [10] et présente des similarités avec les systèmes de classificateurs de Holland. Ainsi, à partir d'une condition d'entrée il permet de déterminer l'action à effectuer. La règle de sélection

d'une action est fondée sur un mécanisme de roulette biaisé. Soit C l'ensemble fini des conditions d'entrée du système de décision et O l'ensemble fini des opérateurs pouvant être sélectionnés. Pour chaque couple c_i, o_j un poids $w_{i,j}$ est défini. Ce poids correspond au potentiel de sélection de l'opérateur o_j pour la condition c_i . La probabilité $P(o_j|c_i)$ de sélectionner l'opérateur o_j dans la condition c_i correspond à la formule 1.

$$P(o_j|c_i) = \frac{w_{i,j}}{\sum_{k=1}^m w_{i,k}} \quad (1)$$

Avec :

$C : (c_i)_{i=1,\dots,n}$; Ensemble des conditions

$O : (o_j)_{j=1,\dots,m}$; Ensemble des opérateurs

$W : (w_{i,j})_{i=1,\dots,n;j=1,\dots,m}$; Matrice de poids

Pour sélectionner un opérateur, le contexte d'optimisation est analysé afin de produire une condition d'entrée du système de décision. Un tirage de type roulette biaisé est ensuite effectué afin de sélectionner l'opérateur. L'intérêt de ce système de décision réside dans sa simplicité de mise en place et de représentation. Grâce à ce système, il est possible de restreindre le choix des opérateurs en fonction du contexte d'optimisation (initialisation du poids à 0) et de favoriser le choix d'un opérateur dans un contexte donné en augmentant le poids correspondant. Ainsi, les mécanismes d'apprentissage auront pour tâche de modifier la matrice des poids de manière à ajuster le choix des opérateurs.

La définition des conditions est un élément essentiel pour le bon fonctionnement du système de décision. Un nombre de conditions trop restreint ne permettra pas de rendre compte du contexte d'optimisation et un nombre trop important de conditions rendra difficile l'initialisation ou l'apprentissage de la matrice des poids. Nous avons choisi, dans un premier temps, de définir une condition pour chaque type de minimum local de la solution courante. Ainsi, la première condition caractérise une solution courante n'étant pas un optimum local, et les autres conditions correspondent aux différentes combinaisons d'optimum local. (Le lecteur pourra se référer à [11] pour une définition des concepts de structure de voisinage et d'optimum local.) Par exemple, si deux structures de voisinages \mathcal{N}_1 et \mathcal{N}_2 sont utilisées lors de la recherche locale, les quatre conditions suivantes seront utilisées par le système de décision : (i) La solution courante n'est pas un optimum local. (ii) La solution courante est un optimum local sur

\mathcal{N}_1 . (iii) La solution courante est un optimum local sur \mathcal{N}_2 . (iv) La solution courante est un optimum local sur \mathcal{N}_1 et \mathcal{N}_2 . Cette définition de l'ensemble des conditions permet de gérer la fréquence ainsi que l'ordre d'application des opérateurs.

À l'initialisation de la matrice des poids, les valeurs $w_{i,j}$ sont initialisées avec un paramètre α . Ainsi, pour une condition donnée, le choix des opérateurs est équiprobable. Cependant, afin d'éviter une application anarchique des opérateurs avant l'apprentissage, plusieurs couples condition/opérateur sont initialisés à la valeur 0. L'initialisation de la matrice des poids correspond aux deux règles suivantes : (i) Un opérateur de diversification ne peut être choisi que si la solution courante est un optimum local sur l'ensemble des structures de voisinage. (ii) Un opérateur de recherche locale fondé sur une structure de voisinage \mathcal{N}_i ne peut être appliqué si la solution courante est un optimum local sur \mathcal{N}_i . Cette initialisation permet d'obtenir un cycle diversification/intensification dans le choix des opérateurs. La phase de diversification correspond à l'application d'un opérateur de diversification, quant à la phase d'intensification, elle correspond à l'application d'une série d'opérateurs d'intensification jusqu'à obtenir une solution courante, optimum local sur l'ensemble des structures de voisinage.

3.5 Mécanismes d'apprentissage

Afin d'adapter le comportement du rôle guide, ici fondé sur le mécanisme de décision précédemment formulé, deux mécanismes d'apprentissage sont utilisés conjointement : apprentissage par renforcement et mimétisme. L'objectif de l'apprentissage dans le cadre de CBM est d'ajuster la matrice de poids du système de décision afin d'améliorer le choix des opérateurs. Ces mécanismes d'apprentissage sont à la base du rôle stratège dans CBM.

Apprentissage par renforcement. L'apprentissage par renforcement consiste à attribuer une sanction aux expériences réalisées afin d'améliorer le comportement. Les points clés de l'apprentissage par renforcement identifiés dans [12] sont la recherche par essais/erreurs et l'affectation retardée de sanctions. Dans le cas de CBM, le rôle guide effectue différentes séquences d'opérateurs grâce au mécanisme de roulette biaisée. Lorsqu'une séquence d'opérateurs a permis d'obtenir un résultat satisfaisant, cette séquence est alors renforcée. Afin de

mettre en place ce mécanisme d'apprentissage par renforcement il est nécessaire de définir : ce qu'est une expérience, les règles de renforcement permettant d'identifier les expériences à renforcer, et les valeurs des sanctions à appliquer aux expériences renforcées.

Par rapport au mécanisme de décision, une expérience correspond à l'application d'un opérateur dans une condition donnée. Ainsi, une expérience est un triplet (condition, opérateur, gain), où le gain est la différence de coût de la solution courante apportée par l'application de l'opérateur. En observant un cycle diversification/intensification, nous considérons que si ce cycle a permis d'améliorer le coût de la meilleure solution trouvée, alors les expériences de gain non nul dans ce cycle sont bénéfiques. Le renforcement correspond à une augmentation des valeurs de poids $w_{i,j}$ associés aux expériences bénéfiques. Deux valeurs de sanctions sont définies σ_1 et σ_2 . La valeur σ_1 est appliquée lorsque, dans le cycle diversification/intensification, seule la valeur de la meilleure solution trouvée est améliorée, et la valeur σ_2 est appliquée si la valeur de la meilleure solution connue est améliorée. Le renforcement d'une expérience correspond à la formule 2.

$$w_{i,j} = w_{i,j} + \sigma \quad (2)$$

Avec :

$(c_i ; o_j)$; Expérience à renforcer
 $w_{i,j}$; Valeur de poids associée à l'expérience
 $\sigma : \{\sigma_1 ; \sigma_2\}$; Valeur de la sanction

Apprentissage par mimétisme. L'apprentissage par renforcement est réalisé individuellement pour chaque instance du rôle stratège. Le mimétisme permet de partager ces connaissances acquises par renforcement [13]. Le principe de l'apprentissage par mimétisme consiste à imiter les comportements efficaces. Dans le cas de CBM, un comportement est identifié comme efficace lorsqu'il a permis d'améliorer la meilleure solution connue.

Lorsque le rôle stratège observe une amélioration de la meilleure solution connue, après avoir effectué un apprentissage par renforcement il soumet la matrice de poids aux autres rôles stratèges afin qu'ils s'en inspirent. Notons W_A la matrice de poids correspondant à la stratégie à imiter, et W_B la matrice de poids de l'imitateur. La matrice de poids résultant de l'imitation correspond à la moyenne pondérée de W_A et W_B selon la formule 3.

$$W_B = (1 - \rho).W_B + \rho.W_A \quad (3)$$

Avec :

W_A ; Matrice de poids à imiter

W_B ; Matrice de poids de l'imitateur

ρ ; Coefficient d'imitation

L'apprentissage par mimétisme est un mécanisme complémentaire à l'apprentissage par renforcement. En effet, dans notre cas le mimétisme permet uniquement de partager des connaissances acquises par renforcement.

4 CBM pour la résolution du problème de tournée de véhicules

Dans cette partie CBM est utilisée pour résoudre le problème de tournées de véhicules avec contrainte de capacité (VRP, *Vehicle Routing Problem*). Cette partie présente la spécialisation de CBM pour traiter le VRP ainsi que les résultats expérimentaux associés.

4.1 Spécialisation de CBM

Le VRP est un problème d'optimisation bien connu dans le domaine du transport. Ce problème a été largement étudié depuis cinq décennies. Il consiste à trouver les routes de moindre coût desservant un ensemble de clients. Nous utiliserons la formulation établie dans [14].

Le VRP est défini sur un graphe $G(V, E)$ où $V = \{v_0, \dots, v_n\}$ est l'ensemble des nœuds et $E = \{(v_i, v_j) / v_i, v_j \in V; i \neq j\}$ représente l'ensemble des arcs. Dans ce graphe, le nœud v_0 correspond au dépôt et les autres nœuds sont les clients. A chaque client $v_i, i \in \{1, \dots, n\}$ est associé une quantité q_i de produit à livrer et un temps d'attente δ_i pour le service. A chaque arc (v_i, v_j) est associé un temps de parcours $c_{i,j}$. Une solution admissible du VRP est un ensemble R de routes respectant les contraintes suivantes : (i) chaque route débute et se termine par le dépôt, (ii) chaque client est desservi par une et une seule route, (iii) la demande total de chaque route ne doit pas dépasser la capacité des véhicules Q , (iv) la durée de chaque route ne doit pas excéder D . L'objectif est de minimiser la durée totale des routes.

Le VRP est un problème NP-Difficile et ne peut être que rarement résolu de manière exacte pour des instances de plus de 100 clients. Plusieurs heuristiques et métaheuristiques ont été proposées pour résoudre le VRP. Une revue de ces méthodes peut être trouvée dans [15]. L'approche CBM semble être adaptée à la résolu-

tion du VRP étant donné qu'il s'agit d'un problème d'optimisation combinatoire difficile sur lequel un grand nombre d'opérateurs peut être défini. La spécialisation de CBM consiste à définir l'ensemble des opérateurs de diversification et d'intensification relatifs aux rôles diversifieur et intensifieur. Pour résoudre le VRP six opérateurs de diversification et six opérateurs d'intensification sont utilisés.

Parmi les opérateurs de diversification deux opérateurs de génération de solutions ont été implantés : Génération gloutonne et génération par rotation (*sweep algorithm*) [16]. Deux procédures de croisement de solution ont également été implantées : Croisement par insertion de route et croisement ordonné [17]. De plus, une procédure de perturbation par suppression/insertion configurée à des taux de 5% et 15% à été utilisée. Les six opérateurs d'intensification correspondent à des procédures de descente locale utilisant les structures de voisinage : *2-opt*, *3-opt*, *1-interchange move*, *1-interchange swap*, *2-String exchange*, *2-String relocation*. Ces structures de voisinage sont décrites dans [18].

4.2 Résultats expérimentaux

L'intérêt de CBM réside dans ces mécanismes d'apprentissage et de coopération. Dans cette partie est mis en évidence le gain apporté par l'apprentissage et la coopération, puis, une comparaison de CBM à d'autres approches est réalisée.

CBM a été implantée en Java et exécutée sur un Pentium 4 à 3GHz avec 1Gb de mémoire vive. Les résultats expérimentaux ont été obtenus sur les 14 instances proposées par Christofides et al. [19]. Ces instances contiennent entre 50 et 199 clients.

Performances de l'apprentissage et de la coopération. Afin de mettre en évidence l'apport de l'apprentissage et de la coopération entre agents, CBM a été testée avec trois configurations différentes et différentes tailles de population. Dans la première configuration de CBM, les agents n'effectuent pas d'apprentissage. Dans la seconde configuration, les agents effectuent uniquement un apprentissage par renforcement. Dans la dernière configuration, les agents réalisent apprentissage par renforcement et mimétisme. Pour chacune des configurations, CBM à été lancé 10 fois sur chaque instance avec un nombre d'agents variant entre 1 et 20. Afin de comparer les différentes configurations,

le nombre total d'applications d'opérateurs a été fixé à 10000 pour l'ensemble de la coalition. Ainsi pour une coalition contenant A agents, un agent effectue en moyenne $10000/A$ applications d'opérateurs. La figure 4 représente les résultats obtenus pour les trois configurations. Chaque valeur correspond à l'écart moyen avec les meilleures solutions connues des instances décrites dans Christofides et al. [19]. Ces meilleures valeurs connues sont issues de [14].

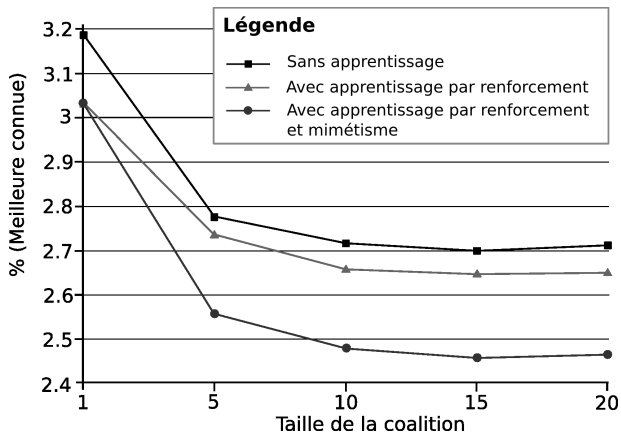


FIG. 4 – Mesure de l'impact de l'apprentissage et de la coopération dans CBM

On remarque sur ce graphique que, quelque soit la taille de la coalition, l'ajout de capacités d'apprentissage aux agents permet d'améliorer le résultat final de l'optimisation. Ensuite, on observe que de meilleurs résultats sont obtenus en augmentant la taille de la coalition. Cela résulte des mécanismes de coopération entre agents. La détérioration des résultats au-delà d'une taille de coalition de 15 agents s'explique par le peu d'application d'opérateurs effectué par les agents individuellement.

Comparaison des résultats de CBM avec d'autres métaheuristiques. Les résultats de CBM ont été comparés avec deux approches de recherche tabou : UTSA (Unified Tabu Search Algorithm) [20] et GTS (Granular Tabu Search) [21]. Le tableau 2 présente les résultats de ces deux approches et de CBM. Les quatre premières colonnes du tableau décrivent les caractéristiques des instances de VRP traités. Une instance de VRP de type C a une contrainte de capacité et une instance de type D a une contrainte de durée. La colonne intitulée "moyenne" pour l'approche CBM correspond à l'écart moyen avec la meilleure solution connue de l'instance. Cette moyenne est obtenue sur dix exécutions. Les deux colonnes suivantes sont respectivement la valeur de l'inter-

valle de confiance à 95% de l'écart moyen, et la durée moyenne d'une exécution de CBM en minutes. Les résultats pour l'approche UTSA et GTS sont tirés de [14]. Les temps d'exécution pour UTSA ont été obtenus sur un Pentium IV à 2GHz et les temps d'exécution de GTS sur un Pentium à 200MHz. En observant les moyennes sur l'ensemble des instances, l'approche CBM semble compétitive vis-à-vis de GTS et d'UTSA, du point de vue des temps d'exécution comme de la qualité des solutions obtenues.

5 Conclusion et perspectives

Cet article introduit AMF (Agent Metaheuristic Framework) un framework pour la modélisation et l'implémentation de métaheuristiques. Ce framework est fondé sur un modèle organisationnel de métaheuristique mettant en évidence quatre rôles au sein des métaheuristiques : intensifieur, diversifieur, guide et stratège. Des lignes méthodologiques ont été définies afin de faciliter la conception ou l'hybridation de métaheuristiques à partir de ce modèle. L'utilisation des approches organisationnelle et multiagent dans le cadre des métaheuristiques encourage la modularité, la réutilisation et facilite la distribution.

En utilisant AMF nous avons conçu une métaheuristique originale nommée CBM (Coalition-Based Metaheuristic) dont le principe est de faire coopérer un ensemble d'agents organisé en coalition. Chaque agent au sein de la coalition dispose des capacités pour résoudre individuellement le problème d'optimisation, cependant la coopération entre agents permet d'améliorer les performances globales. La principale originalité de cette métaheuristique est d'utiliser des mécanismes d'apprentissage permettant d'améliorer la stratégie de recherche au cours de l'optimisation. L'efficacité de CBM a été illustrée en traitant le VRP. Les résultats expérimentaux ont montré que CBM est compétitive vis-à-vis d'autres approches.

Les perspectives de ce travail sont, de compléter le framework AMF afin de tendre vers une méthodologie pour la conception des métaheuristiques. Cela suppose, tout d'abord, de concevoir les outils nécessaires à l'implantation, ensuite, de décrire un processus méthodologique complet. Concernant CBM, il est envisageable d'intégrer de nouveaux mécanismes de décision et d'apprentissage tel que les réseaux de neurones et de l'appliquer à d'autres classes de problèmes.

TAB. 2 – Résultats pour les instances Christofides et al.

Instance	Type	Taille	Meilleure connue	CBM			UTSA		GTS	
				Moyenne (%)	Int. Conf. (95%)	Durée (min.)	Meilleure (%)	Durée (min.)	Meilleure (%)	Durée (min.)
1	C	50	524,61	0,00	±0,00	0,2	0,00	2,32	0,00	0,81
2	C	75	835,26	0,20	±0,22	0,4	0,00	14,78	0,40	2,21
3	C	100	826,14	0,11	±0,11	1,1	0,00	11,67	0,29	2,39
4	C	150	1 028,42	0,75	±0,18	2,1	0,41	26,66	0,47	4,51
5	C	199	1 291,26	1,87	±0,32	7,0	1,90	57,68	2,09	7,50
11	C	120	1 042,11	1,52	±2,00	3,9	3,01	11,67	0,07	3,18
12	C	100	819,56	0,02	±0,04	1,2	0,00	9,02	0,00	1,10
6	C, D	50	555,43	0,00	±0,00	0,4	0,00	3,03	0,00	0,86
7	C, D	75	909,68	0,15	±0,14	0,8	0,00	7,41	1,21	2,75
8	C, D	100	865,94	0,02	±0,03	2,0	0,00	10,93	0,41	2,90
9	C, D	150	1 162,55	1,12	±0,20	6,4	0,46	51,66	0,91	5,67
10	C, D	199	1 395,85	1,41	±0,28	7,2	1,50	106,28	2,86	9,11
13	C, D	120	1 541,14	1,34	±0,19	3,4	0,53	21,00	0,28	9,34
14	C, D	100	866,37	0,00	±0,00	1,6	0,00	10,53	0,00	1,41
Moyenne				0,61	±0,13	2,7	0,56	24,62	0,64	3,84

Références

- [1] S. Voss. *Local Search for Planning and Scheduling*, volume 2148 of *LNCS*, chapter Meta-heuristics : The State of the Art, pages 1–23. 2001.
- [2] C. Blum and A. Roli. Metaheuristics in combinatorial optimization : Overview and conceptual comparison. *ACM Computing Surveys*, 35(3) :268–308, 2003.
- [3] S. Moujahed, N. Gaud, and D. Meignan. A self-organizing and holonic model for optimization in multi-level location problems. In *IEEE Int. Conf. on Industrial Informatics, INDIN 2007*, 2007.
- [4] P. Gruer, V. Hilaire, A. Koukam, and K. Cetnarowicz. A formal framework for multi-agent systems analysis and design. *Expert Systems with Applications*, 23(4) :349–355, 2002.
- [5] E. D. Taillard, L. M. Gambardella, M. Gendreau, and J.-Y. Potvin. Adaptive memory programming : A unified view of metaheuristics. *European Journal of Operational Research*, 135 :1–16, 2001.
- [6] R. Battiti and G. Tecchiolli. The reactive tabu search. *INFORMS Journal on Computing*, 6(2) :126–140, 1994.
- [7] L. Ingber. Adaptive simulated annealing (asa) : Lessons learned. *Journal of Control and Cybernetics*, 25(1) :33–54, 1996.
- [8] R. Hinterding, Z. Michalewicz, and A. E. Eiben. Adaptation in evolutionary computation : a survey. In *IEEE Int. Conf. on Evolutionary Computation*, pages 65–69, 1997.
- [9] H. V. D. Parunak, S. Brueckner, M. Fleischer, and J. Odell. A design taxonomy of multi-agent interactions. *LNCS*, 2935(4) :123–137, 2003.
- [10] S. Ropke and D. Pisinger. An adaptive large neighborhood search heuristic for the pickup and delivery problem with time windows. Technical report, University of Copenhagen, 2005.
- [11] Y. Borenstein and R. Poli. Structure and metaheuristics. In *Genetic And Evolutionary Computation Conference*, pages 1087–1094, 2006.
- [12] R. S. Sutton and A. G. Barto. Reinforcement learning : Introduction. Technical report, Cognitive Science Research Group, 1998.
- [13] T. Yamaguchi, Y. Tanaka, and M. Yachida. Speed up reinforcement learning between two agents with adaptive mimetism. In *IEEE Int. Conf. on Intelligent Robots and Systems*, volume 2, pages 594–600, 1997.
- [14] J. F. Cordeau, M. Gendreau, A. Hertz, G. Laporte, and J. S. Sormany. *Logistics Systems : Design and Optimization*, chapter New Heuristics for the Vehicle Routing Problem, pages 279–297. Springer, 2005.
- [15] M. Gendreau, G. Laporte, and J.-Y. Potvin. *The Vehicle Routing Problem*, chapter Metaheuristics for the capacitated VRP, pages 129–154. 2002.
- [16] A. Wren and A. Holliday. Computer scheduling of vehicles from one or more depots to a number of delivery points. *Operational Research Quarterly*, 23(3) :333–344, 1972.
- [17] C. Prins. A simple and effective evolutionary algorithm for the vehicle routing problem. *Computers and Operations Research*, 31 :1985–2002, 2004.
- [18] G. Laporte and F. Semet. Classical heuristics for the vehicle routing problem. In *Les Cahiers du GERAD*, number G-98-54. GERAD, 1999.
- [19] N. Christofides, A. Mingozzi, and P. Toth. *Combinatorial Optimization*, chapter The vehicle routing problem, pages 315–338. Wiley, 1979.
- [20] J. F. Cordeau, G. Laporte, and A. Mercier. A unified tabu search heuristic for vehicle routing problems with time windows. *Journal of the Operational Research Society*, 52 :928–936, 2001.
- [21] P. Toth and D. Vigo. The granular tabu search and its application to the vehicle routing problem. *INFORMS Journal on Computing*, 15 :333–348, 2003.