



Document d'analyse et de conception

Master 1 ANDROIDE : Arthur Esquerre-Pourtère, Jerome Arjonilla

2020

Robotique en essaim et apprentissage de comportements coopératifs

Introduction

Dans la nature, il est possible d'observer de nombreux comportements coopératifs, dont certains très remarquables. Parmi cette liste de comportements remarquables, on peut citer l'altruisme. En effet, un individu altruiste choisira d'effectuer des actions bénéfiques à d'autres agents et ceci, même si cela peut avoir un impact négatif sur lui-même (référence [1]).

L'altruisme est un comportement qui a déjà été étudié dans de nombreux domaines, et la "Tragedy of Commons" en fait parti, (référence [2]). La "Tragedy of Commons" représente une situation dans laquelle si les individus ne sont pas altruistes, alors ils finiront tous par mourir. Dans cette situation, la quantité de nourriture est limitée et il faut que les individus ne mangent pas plus que nécessaire, afin de permettre à la prochaine génération de survivre.

De nombreuses raisons expliquent l'apparition de l'altruisme, notamment la présence de lien de parenté entre les différents individus, (référence [3]). De plus, des études ont montré que la présence de plusieurs individus altruistes va faire diminuer les déplacements globaux, en effet, ils essayeront de s'entraider entre eux. Ce comportement est ce qu'on appelle la "viscosity" (référence [4]).

L'objectif de ce projet est d'étudier une population d'agents robotiques, d'observer et d'analyser les comportements coopératifs des agents robotiques. Nos agents robotiques seront dans une situation proche de la "Tragedy of commons".

En effet, ils seront dans une arène avec, à leur disposition, des éléments de nourriture. Lorsqu'un agent trouve de la nourriture, il a la possibilité de choisir la quantité de nourriture qu'il veut manger. S'il choisit de manger une grande quantité, alors l'élément de nourriture mettra un temps important avant de réapparaître. Au contraire, si il choisit de manger une faible quantité de nourriture, alors l'élément de nourriture mettra un temps moins important pour réapparaître. Grâce aux choix des agents, on va essayer d'observer des comportements pouvant être interprétés comme de l'altruisme ou de l'égoïsme.

Dans un premier temps, il nous est demandé de refaire les travaux effectués par Nicolas Bredeche sur l'altruisme dans une population de robot (référence [5], [6] et [7]) , en modifiant et en rajoutant des nouvelles fonctionnalités.

En effet, durant ses travaux, Nicolas Bredeche a observé des comportements se rapprochant de l'altruisme. Pour cela, il a effectué un changement brutal dans la difficulté de l'environnement. Ce choc brutal a, dans un premier temps, favorisé les agents égoïstes, mais rapidement, les agents sont devenus plus altruistes. De plus, il a réussi à obtenir des résultats montrant qu'il y avait bien un lien avec le niveau de parenté, et de déplacement des agents.

Et c'est pour cela, qu'il nous est demandé, dans un second temps, d'essayer de

construire des expériences contrôles permettant de mieux observer, et de mieux analyser les comportements qui ont été observés dans ses différents travaux.

Afin de pouvoir faire ceci, nous allons utiliser le simulateur robotique, du nom de "Roborobo3" créé par Nicolas Bredeche (référence [8]).

Méthode

Comme dit précédemment, dans un premier temps, notre travail consiste à reproduire les travaux de recherche effectués par Nicolas Bredeche, en y apportant quelques modifications. Notamment en changeant la forme de l'arène afin qu'elle soit ronde, en augmentant le nombre d'agents présents dans l'arène (de 100 à 200), etc.

De plus, il nous a été demandé d'ajouter des nouvelles informations afin de mieux comprendre les comportements de nos agents robotiques, comme l'élaboration de plusieurs formules de distances.

Deuxièmement, il nous a été demandé de créer des expériences contrôles. Les expériences contrôles sont au nombre de 2, et grâce à celles-ci, on va essayer de comprendre les raisons de la présence de comportements sous optimaux au niveaux des individus et de voir dans quelles conditions elles apparaissent. Les différentes expériences contrôles seront expliquées dans la suite du texte.

Algorithme

Notre travail se base sur l'algorithme mEDEA présenté par Nicolas Bredeche, il s'agit d'un algorithme évolutionniste conçu pour être utilisé sur une population d'agents. Chacun de ces agents a un génome qui influe sur ses actions et son comportement, ainsi qu'un réservoir permettant de stocker les génomes reçus. En effet, à chaque itération, les agents envoient leurs génomes à tout les agents situés à leur portée. À la fin de chaque génération, dans notre cas, au bout de 400 itérations, les agents piochent chacun au hasard dans leur réservoir personnel l'un des génomes reçus, le font muter légèrement, vident leur réservoir personnel, et utilisent le génome ainsi sélectionné pour la prochaine génération.

La principale particularité de cet algorithme évolutionniste est qu'il n'y a pas de fonction objectif clairement définie, contrairement aux autres algorithmes. Toutefois, un génome efficace se répand mieux dans la population, car l'agent qui utilise ce génome vivra plus longtemps et partagera ce génome avec plus d'agents.

Une autre particularité de l'algorithme mEDEA est la façon dont les agents peuvent naître à la génération suivante. En effet, il existe différentes situation pour renaître à la prochaine génération.

Dans un premier temps, pour renaître, un agent doit forcément avoir reçu un génome d'un autre agent. Dans un second temps, si un agent réussit à avoir de l'énergie jusqu'à la fin de la génération, alors il renaîtra avec l'énergie restante, et au contraire, si sa jauge d'énergie tombe à zéro pendant la génération, alors il renaîtra avec l'énergie initial.

Les caractéristiques de mEDEA sont très intéressantes, car elles forcent les individus à prendre de la nourriture pour survivre et à partager leur génome mais, en même temps, les individus doivent apprendre à ne pas prendre trop de nourriture afin de ne pas commencer la prochaine génération avec une quantité moins importante que les autres individus.

Configuration expérimentale

Afin d'obtenir des résultats cohérents et fiables d'un point de vue statistique, chaque couple de paramètres a été évalué 20 fois. Il aurait été mieux de faire plus d'évaluations à chaque fois mais la puissance de nos ordinateurs nous a limité.

Global Parameter	Value
arena	circular 1000*1000 pixels
population size	200 agents
number of energy points	800
selection	variable (See text)
number of iteration	800 000
lifetime (i.e. generation duration)	400 iterations

Agent Parameter	Value
agent maximum energy level	800
agent initial energy level	200
agent energy consumption	1 energy unit per step
amount of food taken by an agent	fixed or dynamic (see text)
proximity sensor range	32
agent maximum rotational velocity	30 deg/step
agent translational velocity	2 deg/step
variation operator	gaussian mutation with σ parameter

Food Parameter	Value
energy per food item	100
food item regrow delay	between 25 and 4000 steps (see text)
food item diameter	2
food item footprint diameter	4

Figure 1: Paramètres pour les expérimentations

Dans le tableau ci-dessus, on peut voir les paramètres que nous avons choisis pour effectuer nos observations. Une grande partie des paramètres ont été choisis

en utilisant les paramètres déjà utilisés dans les précédents articles de Nicolas Bredeche et d'autres ont été choisis après discussions et réflexions toujours avec Nicolas Bredeche.

Comme dit précédemment, l'un des objectifs de notre sujet est d'essayer d'observer des comportements se rapprochant de l'altruisme et/ou de l'égoïsme. Afin d'essayer d'obtenir cela, nous allons faire varier certains paramètres.

Les paramètres en question sont :

- TF = Time food = Food item regrow delay
- Q = Quantity = amount of food taken by an agent

TF

La variable TF correspond aux temps maximaux pour que la nourriture réapparaisse. Les valeurs possibles pour cette variable sont :

- TF = 25
- TF = 100
- TF = 400
- TF = 2000
- TF = 4000

Lorsque la valeur de la variable TF est élevée, on dit que l'environnement est très hostile. En effet, si le temps de réapparition de la nourriture est de 4000, cela signifie que dans la pire des cas, il faudra 10 générations pour que cette nourriture réapparaisse.

Afin d'arriver à des valeurs de TF importantes, nous avons décidé d'augmenter les valeurs de TF petit à petit après chaque génération. Ainsi, après chaque génération, la valeur de TF augmente de 10 et ce jusqu'à arriver à la valeur de TF souhaitée.

Lorsque la valeur de TF augmente au fur et à mesure, on dit que l'environnement est stressant pour les agents.

Q

La variable Q représente le pourcentage de quantité de nourriture que va choisir l'agent. Dans certains cas de figure, nous avons décidé de fixer cette valeur, et dans d'autres cas, c'est l'agent qui va décider dynamiquement de cette valeur.

Lorsque Q est fixe, les valeurs possibles sont :

- Q = 0.7
- Q = 1

Expérience de base

Le rôle de l'expérience de base est de nous permettre d'analyser l'impact des modifications apportées par les expériences contrôles, et ainsi d'essayer d'analyser les comportements coopératifs.

Dans cette expérience, les agents posséderont un réservoir individuel et nous utiliserons l'algorithme mEDEA.

Expérience Contrôle 1

L'objectif de l'expérience contrôle 1 est d'étudier l'influence de la dispersion d'une population sur l'apparition des comportements pouvant être interprété comme de l'altruisme entre les individus.

Cette expérience se base sur ce que nous avons fait auparavant, mais nous avons changé le fonctionnement du réservoir à génome. Ainsi, au lieu d'avoir un réservoir personnel dans lequel ils piochent au hasard un génome, les agents vont partager un réservoir global. C'est à dire, lorsqu'un agent reçoit un génome, celui-ci va aller dans le réservoir global, et à la fin de la génération, l'agent va tirer au hasard dans le réservoir global un nombre de génomes égale au nombre de génomes qu'il a reçus, et ensuite il utilisera pour la génération suivante, encore une fois au hasard, l'un des génomes qu'il a pioché, après l'avoir fait muter.

Ainsi, cela aura pour effet de disperser les génomes dans notre environnement. Le but est d'observer si l'apparition des comportements de partage de ressources est due au fait que la localité des échanges favorise le partage de ressources entre des familles d'individus.

Expérience Contrôle 2

Nous avons aussi fait une deuxième expérience contrôle dans laquelle nous avons une approche très différente. En effet, dans cette expérience, nous utilisons K arènes et pour chacune de ces arènes, tous les individus qui la composent sont des clones, qui ont le même génome. Les k arènes sont en compétition les unes contre les autres, et à la fin d'une génération, on sélectionne la meilleure arène.

La sélection de la meilleur arène est décidée en fonction du nombre d'individus encore en vie. Ainsi, l'arène avec le plus d'individus à la fin est l'arène qui sera utilisée pour la sélection.

À la génération suivante, on assignera à chaque arène une version mutée du génome pris de la meilleure arène, et tout les agents d'une arène prendront le génome affecté à celle-ci.

Comme cette expérience contrôle est très différente de l'expérience contrôle 1

et de l'expérience de base, nous ne pouvons, malheureusement, pas tester les mêmes paramètres. En effet, dans les expériences précédentes, nous avons pu effectuer des expériences sur des valeurs de TF allant jusqu'à 4000. Cependant, dans nos expériences, nous nous sommes donné à chaque fois un budget de 800000 itérations. Ici, nous ne pouvons aller à des TF aussi élevés, car il faut K fois plus d'itérations pour augmenter le TF. Ainsi, dans cette expérience contrôle nous avons seulement pu tester pour des valeurs de TF allant jusqu'à 400.

Contrairement à l'expérience contrôle 1 qui a pour objectif de disperser les génomes sur la carte, cette expérience a pour effet d'homogénéiser la population. On s'attend donc à ce que ces individus tendent à être plus altruistes.

Contrôle de la dispersion spatial

La dispersion spatial est un facteur très important pour nous. En effet, les précédents travaux ont montré que la dispersion spatiale possède un impact sur la stratégie de consommation des agents, et sur un potentiel altruisme/égoïsme. En effet, une valeur faible de dispersion spatiale, peut à la fois favoriser l'exploitation des mêmes éléments de nourriture et augmenter la kin selection, et ainsi avoir un impact sur le niveau d'altruisme.

Et c'est pour cela que le contrôle de la distance parcourue est important pour nous. La distance permet en effet de donner un indicateur pour savoir si les agents ont tendance à explorer loin de leur position initiale, ou pas. Il est donc important de pouvoir la monitorer du mieux possible. Pour cela, nous avons modifié certaines formules de calcul de la distance et en avons rajouté d'autres. Ces formules permettent de mesurer la distance moyenne parcourue par nos agents, à chaque génération.

Ainsi, nous avons créé 3 formules de distance qui représentent :

- D1 : Distance 1
 - Formule permettant d'observer la distance totale parcourue par les agents, en prenant un pas de 5 itérations.
- D2 : Distance 2
 - Formule permettant d'observer la différence de position entre la première et la dernière itération.
- D3 : Distance 3
 - Formule permettant de découper la carte en cases et de mesurer le nombre de cases sur lesquels un agent est passé.

Pour l'explication des formules, nous avons N ="nombre d'agent", $x(t)$ ="Position en x au temps t", $y(t)$ ="Position en y au temps t". La valeur minimale

de t est 1, la valeur maximale est 400 dans nos expériences, car nos générations durent 400 steps.

Formule D1

La formule D1 est égale à

$$\frac{1}{N} * \sum_{i=1}^N \sum_{t \% 5 = 0} \sqrt{(x(t) - x(t-5))^2 + (y(t) - y(t-5))^2}$$

Cette formule correspond donc la somme des distances parcourues toutes les 5 itérations. Grâce à cette formule, nous espérons observer si, localement, il y a des conditions/expériences qui favorisent ou défavorisent les déplacements des agents. En effet, une hypothèse est que si nos agents n'ont pas besoin de se déplacer pour chercher de la nourriture (car il y en a en grande quantité), alors il y aura plus de chance d'observer une valeur faible. Et qu'au contraire, si nos agents ont besoin de se déplacer pour chercher de la nourriture (car il y en a peu), alors il y aura plus de chance d'observer une valeur importante.

Formule D2

La formule D2 est égale à

$$\frac{1}{N} * \sum_{i=1}^N \sqrt{((x(400) - x(1))^2 + (y(400) - y(1))^2)}$$

Cette formule permet donc d'observer d'observer la différence de position entre l'itération $t=1$ et l'itération $t=400$. Grâce à ceci, nous souhaitons observer les déplacements globaux de nos agents, et observer si nos agents restent proches de leurs point de départ, ou au contraire, si ils se sont déplacés.

Formule D3

La formule D3 permet aussi de déterminer les déplacements globaux de nos agents, mais d'une façon différente.

En effet, pour cette formule, nous avons du dans un premier temps diviser notre carte en carrés. Les carrés sont de taille $4*4$ pixels et nous avons au total 62 500 carrés sur notre carte.

Si un agent explore un carré, nous incrémentons une valeur de +1 pour cette case et à la fin de la génération, nous déterminons la moyenne du nombre de carrés explorés par chaque agent divisé par le nombre total de carrés. Ainsi, grâce à cette variable, on peut voir avec précision les déplacements globaux de nos agents.

En revanche, il est important de signaler que notre carte est circulaire et que les agents n'ont pas accès à tous les carrés.

Contrôle de la stratégie de consommation

La réapparition de la nourriture est impactée par plusieurs facteurs, en effet, comme on peut voir dans la formule ci-dessous :

$$TF_{actual} = \frac{TF * E_{recolte}}{E_{maxEnergie}}$$

avec TF la variable Time food qui est définie dans la partie "Configuration expérimentale", $E_{recolte}$ qui représente la quantité récoltée par un agent, $E_{maxEnergie}$ qui représente l'énergie d'un item de nourriture (Ici 100).

En fonction de nos paramètres définis précédemment, on peut estimer des valeurs théoriques sur le nombre de nourriture disponible pour chaque agent, et ceci en fonction de la difficulté de l'environnement (TF).

Dans notre arène il y a 200 agents avec chacun une énergie initiale de 200, de plus, chaque déplacement coûte 1 point d'énergie, ainsi, un agent doit pouvoir dépenser 400 points d'énergie, ou plus, pour survivre jusqu'à la fin d'une génération. Autrement dit, cela signifie qu'il faut au moins 2 items de nourriture pour qu'un agent survive durant toute sa génération. En effet, chaque item d'énergie peut donner au maximum 100 ($E_{maxEnergie}$). De plus, comme il y a un nombre d'items de nourriture de 800, cela signifie qu'il faut au moins avoir ($NombreItemEnergiePourSurvivre * NombreAgent$) $2*200 = 400$ item d'énergie pour que toute la population survive.

Ainsi, en fonction du TF, on peut réussir à déterminer les valeurs théoriques. Par exemple, dans la situation où l'environnement possède un TF de 400, cela signifie que dans le pire des cas, la nourriture mettra un temps de $\frac{400*100}{100} = 400$ afin de réapparaître.

Autrement dit, cela signifie que pour un TF de 400, il faut une génération afin que la nourriture réapparaisse, et donc que durant une génération, il y aura en moyenne 800 items d'énergie. Dans cette situation, cela signifie qu'il y a en moyenne 4 item d'énergie pour chaque agent. Dans ces conditions, cela signifie que toute la population peut survivre.

En revanche, en augmentant le TF, on arrivera à un niveau tel qu'il n'y aura pas assez d'items d'énergie pour toute la population. De façon précise, ce niveau arrive lorsque le TF est de 800, ainsi, si on observe une valeur de TF d'au moins 801, cela signifie qu'au moins un agent ne pourra pas survivre.

Comme précisé dans nos paramètres, notre TF peut prendre les valeurs de 25, 100, 400, 2000 et 4000. Dans les situations où le TF est de 2000 ou 4000, alors il y aura forcément des agents qui ne pourront survivre. Plus précisément, si le TF est de 2000, cela signifie qu'il faut 5 générations pour qu'un item d'énergie réapparaisse, ainsi en moyenne, par génération, il y a 160 items d'énergie, or cela signifie qu'au maximum, il y aura 80 individus qui pourront survivre. Et pareillement, lorsque le TF vaut 4000, cela signifie, qu'au maximum seulement 40 individus pourront survivre.

Bien sûr, cette valeur est théorique, et même si, il y a assez d'énergie pour que toute la population survive, il est fort possible qu'une partie n'arrive pas à survivre. En effet, la nourriture peut être très loin pour agent, il peut y avoir de l'égoïsme, etc. De plus, cette valeur n'est valable que pour une seule génération, si des agents survivent, ils réapparaîtront à la génération suivante avec surement moins de points d'énergie et auront donc besoin de plus de ressources pour survivre.

Résultat et Analyse

Notre partie résultat sera composée de 2 grands axes. En premier lieu, nous effectuerons et analyserons l'expérience de base.

Deuxièmement, nous étudierons les expériences contrôles en les comparant avec nos résultats obtenus dans la première partie. En effet, les résultats obtenus dans la première partie nous serviront de résultats témoins, et nous permettront de mieux observer l'impact des différentes expériences contrôles.

Expérience de base

Comme précisé dans la partie "Expérience de base" de "Méthode", pour l'expérience de base, nous avons utilisé les paramètres de bases, l'algorithme mEDEA et des réservoirs individuels pour chaque agent.

Nombre d'agents vivants

Nous avons expliqué dans la partie "Contrôle de la stratégie de consommation", qu'il y a assez de nourriture pour permettre à tous nos agents de survivre et ceci jusqu'à un TF 800. Comme on peut voir dans la figure [2], on observe que nos agents ne survivent pas tous, et ceci, même si la difficulté de l'environnement n'est pas très importante (TF 25 et 100).

Ce résultat est cohérent, car dans la partie "Contrôle de la stratégie de consommation", les hypothèses étaient très fortes. En effet, les items de nourritures peuvent se trouver loin d'un agent.

De plus, comme précisé dans la partie "Algorithme", mEDEA est fait de tel sorte qu'il n'est pas obligatoire de survivre durant toute une génération pour revivre à la prochaine, ainsi, c'est possible que cela soit une stratégie des agents afin d'essayer de ne pas avoir le maximum d'énergie.

De façon plus précise, si on observe chaque TF, on peut voir des informations importantes, voir figures [3], [4], [5], [6] et [7] en annexe.

Premièrement, on observe que pour chaque TF, lorsque les agents sont les plus gourmands (prennent le plus de quantité possible), alors il y aura à la fin plus

de survivants. De plus, lorsque l'environnement devient très hostile (TF 2000 et 4000), l'écart type se réduit fortement.

De plus, il est intéressant d'observer que lorsque la quantité est dynamique, et que l'environnement n'est pas très hostile (TF 25, 100 et 400) le nombre de survivants est compris entre le nombre de survivants lorsque que la quantité fixe est de 70 et le nombre de survivants lorsque que la quantité fixe vaut 100. En revanche, lorsque l'environnement devient plus hostile (TF 2000 et 4000), on observe que le nombre de survivants est bien plus faible que lorsque les quantités étaient fixe. On observe bien cette tendance, notamment lorsque le TF est à 2000, car en quantité dynamique, on observe un nombre de survivant vers 17, alors que lorsque la quantité fixe est de 70 on observe un nombre de survivant proche de 25, et lorsque la quantité fixe est de 100, alors le nombre de survivants est proche de 35.

Quantité

Un autre résultat intéressant est le choix de quantité de nourriture, lorsque celle ci est dynamique, voir figure [8] et [9].

Comme, on peut le voir, les agents ont tendance à prendre une quantité de nourriture plutôt importante (Proche de 90). Cependant, on arrive à observer une différence en fonction de la difficulté de l'environnement. En effet, on observe que lorsque le TF est faible (TF 25 et 100), alors la quantité prise par les agents est d'environ 87, au contraire, lorsque le TF augmente (TF 4000), on observe que la quantité se rapproche de 92.

En revanche, d'après nos résultats, on observe un écart type important, ainsi, nous ne pouvons pas affirmer qu'il y ait une réelle différence en fonction de la difficulté de l'environnement. Cependant, il est important de rappeler que nous n'avons pu faire nos expériences seulement sur 20 runs différents, ainsi, il est possible, que cet écart soit réellement significatif mais il faudrait faire un nombre de run plus important pour en être sûr.

Expériences contrôle

Nous allons maintenant comparer nos expériences témoins avec les résultats obtenus lors de la première partie. Ceci afin d'observer si les changements que nous avons apportés lors de ces expériences ont eu un impact sur les comportements coopératifs et sur la survie de nos agents.

Expérience contrôle 1 : réservoir global

Nous allons donc commencer par comparer les résultats de l'expérience de base avec ceux du de l'expérience contrôle 1. Pour rappel, il s'agit de l'expérience où les agents partagent entre eux un réservoir global.

D’une manière générale, les résultats obtenus avec cette méthode sont très similaires à ceux obtenus pour l’expérience des réservoirs individuels, et ce quel que soit la valeur du TF, qu’il s’agisse d’une quantité fixe, ou non.

Nombre d’agents vivants

En effet, comme on le voit en comparant les figures [10] et [11], les nombres d’agents encore en vie à la fin d’une génération sont très similaires entre nos deux expériences. Nos résultats semblent donc montrer que malgré la dispersion génétique cela n’a pas eu d’influence sur la capacité de la population à survivre.

Distance parcourue

De plus, les graphes [12] et [13] nous montrent que dans les 2 expériences, les agents semblent parcourir des distances similaires à chaque génération, ce qui laisse penser que le changement de fonctionnement du réservoir n’encourage pas nos agents à plus ou moins se déplacer. On observe toutefois pour certaines valeurs de TF que les agents dans l’expérience du réservoir individuel semblent se déplacer un peu plus que ceux qui utilisent un répertoire global. Notamment si on observe la métrique D2. On le voit par exemple sur la figure [24] .

Toutefois, cette différence semble assez faible au vu des écarts-type observés pour cette métrique, il faudrait donc faire tourner nos expériences plus de 20 fois chacune afin de voir si cette différence persiste, ou s’il ne s’agit que de bruit.

Quantité

Enfin, lorsque la quantité est dynamique, ces deux approches semblent donner des résultats similaires quant à la quantité de nourriture prise par les agents. En effet, on le voit bien sur le graphe [18] par exemple. Les données semblent suivre la même courbe dans les deux cas, et on n’observe pas de différence autre que du bruit. On observe ceci pour tout les TF sauf lorsque $TF=2000$, mais étant donné l’écart-type, il ne s’agit probablement la aussi que de bruit. (voir [14] et [15]). Cela semble signifier que le fait de mélanger les agents et leurs génomes dans l’arène n’a pas d’influence sur l’apparition de comportements altruistes dans une population.

Conclusion expérience contrôle 1

Ces résultats sont toutefois à prendre avec précaution. En effet, bien que notre méthode permette de disperser les génomes dans l’arène, une hypothèse plausible est que l’arène est en fait trop petite pour qu’une dispersion efficace ai lieu. Pour confirmer et infirmer cela, il faudrait refaire nos expériences dans une arène beaucoup plus grande avec plus d’agents. Et peut-être que dans cette situation, des familles différentes d’individus se créeraient dans l’expérience de base, et que notre expérience contrôle 1 permettrait de disperser ces familles, donnant des résultats potentiellement différents.

C’est un axe de recherche futur qui est peut-être prometteur, mais nous n’avons pas pu le faire pour des raisons techniques (les calculs sont très longs à faire sur nos ordinateurs, cela multiplierais le temps de calcul).

Expérience contrôle 2 : population de clones

Dans cette sous section, nous allons étudier les résultats obtenus pour l'expérience contrôle 2 et les comparer avec notre expérience de base. Pour rappel, il s'agit de l'expérience avec les K arènes différentes, et dans cette expérience contrôle, les valeurs de TF possibles sont 25, 100 et 400.

De façon globale, il est important de noter que la grande majorité des résultats disponibles pour l'expérience contrôle 2, ont tous un écart type important. Malgré ceci, nous avons réussi à observer des tendances.

Quantité

Une tendance observable concerne la quantité dynamique choisie par les agents. En effet, on observe que la quantité de nourriture choisie par les agents lors de l'expérience contrôle 2 est nettement moins importante que la quantité choisie lors que l'expérience de base, voir figure [16], [17] et [18].

En comparant les résultats, on observe que pour l'expérience de base, la quantité de nourriture était d'environ 90. Et au contraire, dans cette expérience contrôle, on observe que la quantité de nourriture est en moyenne proche de 80, et ceci pour chaque TF disponible.

Les caractéristiques de l'expérience contrôle 2, nous ont permis de faire diminuer de façon importante la quantité de nourriture prise, cela semble indiquer que, la présence d'un unique gène dans la population permet de ne pas favoriser une consommation de nourriture plus importante que nécessaire.

Nombre d'agents vivants

En étudiant la variable correspondant aux nombres d'agents vivants, alors on observe 2 tendances différentes. Les résultats sont identiques lorsque la quantité est fixe et lorsque la quantité est dynamique. Par exemple, on peut voir les deux tendances sur les figures [19], [20] et [21] qui correspondent à la situation où la quantité est dynamique.

En effet, lorsque le TF est faible (25 et 100), on observe que le nombre d'agents vivants est plus important pour l'expérience contrôle 2. Par exemple, dans la situation avec un TF de 25, il y a environ 180 agents qui sont vivants, or il y a environ 150/160 agents en vie dans l'expérience de base.

Au contraire, on observe le comportement opposé lorsque l'environnement devient plus hostile (TF 400). En effet, dans cette situation, on observe que le nombre d'agents est moins important dans l'expérience contrôle 2. En comparant le graphique, on peut voir que pour l'expérience contrôle 2, il y a environ 100 agents en vie, alors que pour l'expérience de base, il y a environ 120 agents en vie.

Finalement, on peut dire que lorsque l'environnement n'est pas hostile, les

agents s'en sortent mieux dans l'expérience contrôle 2, et qu'au contraire, lorsque l'environnement devient plus hostile alors les agents survivent mieux dans l'expérience de base.

Malgré la présence d'une population qui a pour objectif de maximiser le nombre de survivant, cela n'a pas permis d'obtenir plus d'individus en vie que pour l'expérience de base lorsque l'environnement devient trop hostile. Il est possible que cela soit dû à l'écart type trop important, et le manque de runs qui ne permet pas d'obtenir des résultats assez fiable, ou alors, cela signifie que leur stratégie n'est pas assez efficace.

Distance parcourue

Ici, on observe que toutes les formules de distances sont très intéressante. En effet, pour toutes les formules de distance (D1,D2 et D3), on observe que les agents se déplacent beaucoup moins dans l'expérience contrôle 2. Ce résultat est observable pour toutes les valeurs de quantité (fixe ou dynamique).

En regardant de plus près, on observe que la différence est plus importante pour la formule D2 (Distance correspondant à l'écart entre le point de départ et le point final). En observant la figure [23], on passe d'une valeur d'environ 200 (Expérience de base) à une valeur proche des 50 (Expérience contrôle 2).

De plus, comme on peut voir dans les figures [22], [23] et [24], on observe que la variable distance D2 augmente en fonction de la difficulté de l'environnement. En effet, on peut voir que lorsque l'environnement possède un TF de 25, on observe une valeur moyenne inférieure à 50, et lorsque l'environnement devient plus hostile (TF 400), on observe une valeur moyenne proche de 90/100. Ce résultat est probablement dû à la nécessité des agents d'avoir un certain niveau de nourriture, et donc de faire une plus grande distance pour en trouver.

En revanche, ce résultat est seulement observable pour la variable D2, dans les autres formules de distance, il n'y a pas d'augmentation en fonction de la difficulté de l'environnement.

Conclusion expérience contrôle 2

Toutes les informations décrites ci-dessus nous montrent que l'expérience contrôle 2 a favorisé l'altruisme. En effet, comme expliqué dans l'introduction, la littérature a observé qu'il y a un lien négatif entre la dispersion spatiale des individus et l'altruisme. En effet, une dispersion spatiale importante est souvent liée à des individus moins altruistes. Et au contraire, lorsque la dispersion spatiale est plus faible, on observe une présence plus élevée d'individus altruiste.

Dans notre expérience, il n'y a pas de dispersion des gènes puisque tout les individus ont le même génome dans chaque arène, de plus les individus ont probablement intérêt à être altruistes puisque pour faire perdurer leur génome,

ils partagent un objectif commun : avoir le plus d'individus en vie à la fin de la génération. Et c'est exactement ce que nous observons.

Le fait que la distance parcourue par les agents soit inférieure à celle parcourue lors de l'expérience de base s'explique par le fait que, dans cette situation, les agents n'ont plus besoin de rencontrer d'autres agents afin de partager leur génome.

Conclusion et perspective

Les résultats obtenus jusqu'ici nous ont permis d'observer des phénomènes intéressants. En particulier concernant l'apparition de comportements pouvant être interprétés comme de l'altruisme dans le cas de l'expérience contrôle 2. L'hypothèse est que l'apparition de ces comportements est due premièrement au fait que les agents partagent tous les mêmes gènes et deuxièmement car ils partagent l'objectif commun d'avoir le plus d'agents en vie possible à la fin d'une génération.

Concernant l'expérience contrôle 1, la valeur de la quantité obtenue lors de nos expérimentations était exactement la même que lors de l'expérience de base, ce qui semble aller à l'opposé de l'hypothèse comme quoi l'altruisme serait moins présent lorsque la dispersion augmente.

Par conséquent, de futurs travaux pourraient permettre de mieux comprendre nos résultats ou bien de confirmer/infirmier nos hypothèses. En effet, une hypothèse, déjà expliquée plus haut, est que la population se mélange trop facilement car notre arène est trop petite. Il pourrait être intéressant de faire la même expérience avec une arène plus grande.

De plus, la puissance de calcul de nos ordinateurs étant limitée, nous n'avons fait tourner chaque configuration que 20 fois à chaque fois, faire plus de replicats pourrait permettre d'obtenir des résultats plus fiables statistiquement.

D'autres fonctionnalités intéressantes pourraient être implémentées afin de voir les effets qu'elles auraient sur les comportements altruistes. Notamment, on pourrait faire dépendre la consommation d'énergie des déplacements, ce qui pourrait pousser nos agents à faire un compromis entre ce disperser sur la carte en ce déplaçant ou bien moins ce déplacer afin d'économiser de l'énergie.

On pourrait aussi changer la fonction de réapparition de la nourriture afin de la rendre non linéaire, on peut voir des exemples de fonctions potentiellement intéressantes sur la figure [25]. Pour calculer le temps de réapparition on utiliserais la valeur de y ainsi : $tempsdereapparition = regrowTimeMax * y$

La fonction f dans le graphe correspond à la fonction suivante :

$$\begin{aligned} f(x) &= (1/2) * ((2 * x)^n) \text{ si } x \leq 0.5 \\ f(x) &= 1 - ((1/2) * ((2 * (1 - x))^n)) \text{ si } x > 0.5 \end{aligned}$$

Enfin, une dernière idée serait que lorsqu'une ressource réapparaît, celle-ci a

une quantité de points d'énergie assez faible et que la ressource a besoin d'un certain laps de temps pour arriver à maturation. Dans cette situation, les agents pourraient détecter le niveau de maturation d'une ressource et décider, ou non, de la manger. Ainsi ils devraient faire un compromis entre prendre la ressource ou bien ne pas la prendre tant que celle-ci n'a pas plus d'énergie. On suppose qu'un agent altruiste aurait tendance à attendre que la nourriture ait un niveau de maturation élevé, car attendre cela serait potentiellement bénéfique pour les agents voisins de celui-ci.

Remerciements

Nous tenons à remercier toutes les personnes qui ont contribué directement ou indirectement à la réalisation de ce projet.

En particulier nous souhaitons remercier M.Bredeche, notre encadrant de projet qui nous a aidé en nous donnant des pistes de réflexion ainsi que des conseils pour résoudre les problèmes auxquels nous avons été confrontés.

Nous remercions également toute l'équipe enseignante du master informatique de Sorbonne Université, qui nous a aidé à obtenir les connaissances théoriques et pratiques nécessaires à la réalisation de ce type de projet.

Références bibliographiques

- [1] Stuart A. West, Andy Gardner, and Ashleigh S. Griffin. 2006. Altruism. English. *Current Biology*, 16, 13, (July 2006), R482–R483. Publisher: Elsevier. DOI: 10.1016/j.cub.2006.06.014.
- [2] Garrett Hardin. 1968. The Tragedy of the Commons. en. *Science*, 162, 3859, (December 1968), 1243–1248. Publisher: American Association for the Advancement of Science. DOI: 10.1126/science.162.3859.1243.
- [3] David C. Queller and Joan E. Strassmann. 2002. Kin selection. English. *Current Biology*, 12, 24, (December 2002), R832. Publisher: Elsevier. DOI: 10.1016/S0960-9822(02)01344-1.
- [4] JOSHUA Mitteldorf and DAVID SLOAN Wilson. 2000. Population Viscosity and the Evolution of Altruism. *Journal of Theoretical Biology*, 204, 4, (June 2000), 481–496. DOI: 10.1006/jtbi.2000.2007.
- [5] Nicolas Bredeche and Jean-Marc Montanier. 2010. Environment-Driven Embodied Evolution in a Population of Autonomous Agents. In *Parallel Problem Solving from Nature, PPSN XI*. Robert Schaefer, Carlos Cotta, Joanna Kolodziej, and Günter Rudolph, editors. Springer Berlin Heidelberg, Berlin, Heidelberg, 290–299. DOI: 10.1007/978-3-642-15871-1_30.
- [6] NTNU, Jean-Marc Montanier, and Nicolas Bredeche. 2013. Evolution of Altruism and Spatial Dispersion: an Artificial Evolutionary Ecology Approach. In *Advances in Artificial Life, ECAL 2013*. MIT Press, (September 2013), 260–267. DOI: 10.7551/978-0-262-31709-2-ch040.

- [7] Jean-Marc Montanier and Nicolas Bredeche. 2011. Surviving the Tragedy of Commons : Emergence of Altruism in a Population of Evolving Autonomous Agents. *European Conference on Artificial Life*, (August 2011).
- [8] Nicolas Bredeche, Jean-Marc Montanier, Berend Weel, and Evert Haasdijk. 2013. Roborobo! a Fast Robot Simulator for Swarm and Collective Robotics. *arXiv:1304.2888 [cs]*, (April 2013). arXiv: 1304.2888.
- [9] Stuart J. Russell. 2010. *Artificial intelligence : a modern approach* /. Prentice Hall.
- [10] Richard Watson, Sevan Ficici, and Jordan Pollack. 2002. Embodied Evolution: Distributing an Evolutionary Algorithm in a Population of Robots. *Robotics and Autonomous Systems*, 39, (April 2002), 1–18. DOI: 10.1016/S0921-8890(02)00170-7.
- [11] Joseph Early. 2020. What’s hot in Multi-Agent Systems? en. (February 2020). Retrieved 03/15/2020 from <https://medium.com/swlh/whats-hot-in-multi-agent-systems-4b0f348e68bd>.
- [12] Bastien Verdebout. 2020. Machine learning : from idea to reality. en-GB. (March 2020). Retrieved 03/15/2020 from <https://www.ovh.com/blog/machine-learning-from-idea-to-reality/>.
- [13] Richard Dawkins. 2016. *The Selfish Gene*. en. Oxford University Press, (June 2016).
- [14] Léo Cazenille, Nicolas Bredeche, Heiko Hamann, and Jürgen Stradner. 2012. Impact of Neuron Models and Network Structure on Evolving Modular Robot Neural Network Controllers. en. In *GECCO - Genetic and Evolutionary Computation Conference*, 89.
- [15] Ridha El Hamdi, Mohamed Njah, and Mohamed Chtourou. 2008. Multi-layer perceptron training using an evolutionary algorithm. *International Journal of Modelling, Identification and Control*, 5, 4, (January 2008), 305–312. Publisher: Inderscience Publishers. DOI: 10.1504/IJMIC.2008.023515.
- [16] Robert Axelrod and William D. Hamilton. 1981. The Evolution of Cooperation. *Science*, 211, 4489, 1390–1396. Publisher: American Association for the Advancement of Science.

Annexe

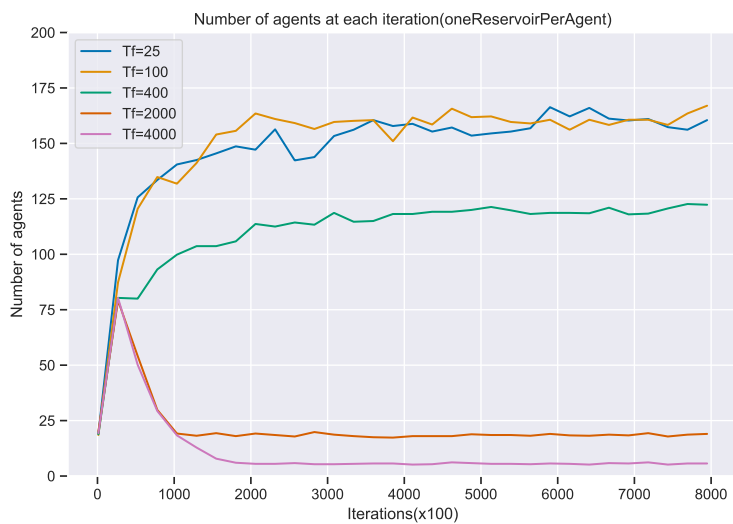


Figure 2: Nombre d'agents après chaque itération pour l'expérience de base

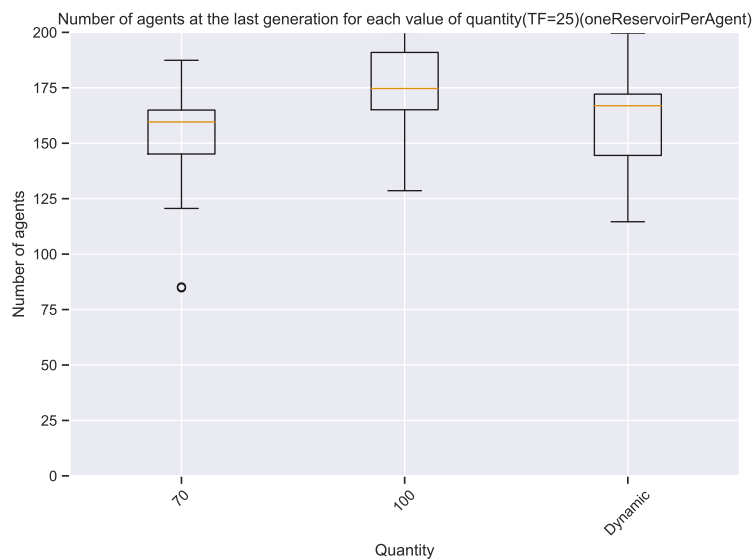


Figure 3: Nombre d'agents vivants à la dernière génération en fonction de la quantité prise et lorsque l'environnement possède un TF de 25

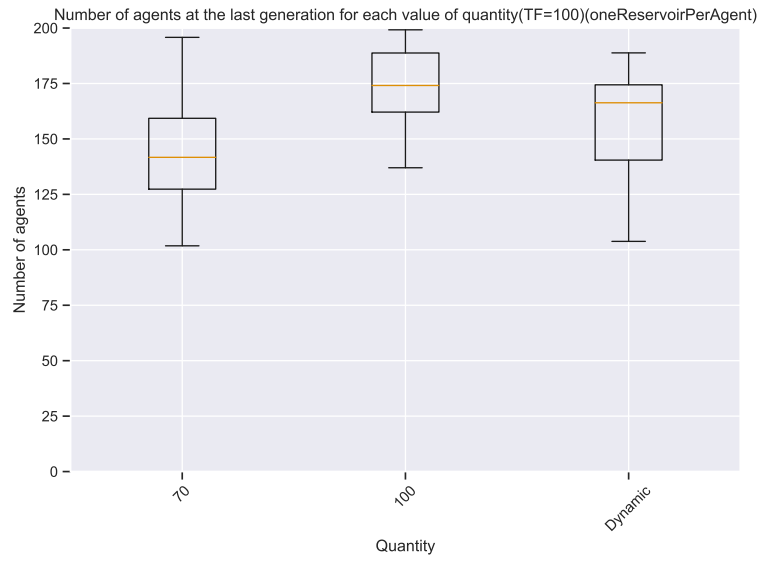


Figure 4: Nombre d'agents vivants à la dernière génération en fonction de la quantité prise et lorsque l'environnement possède un TF de 100

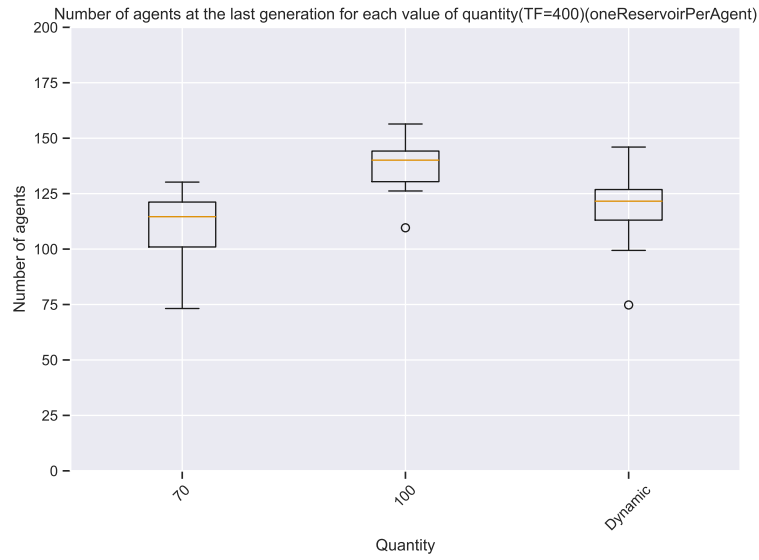


Figure 5: Nombre d'agents vivants sur la dernière génération en fonction de la quantité prise et lorsque l'environnement possède un TF de 400

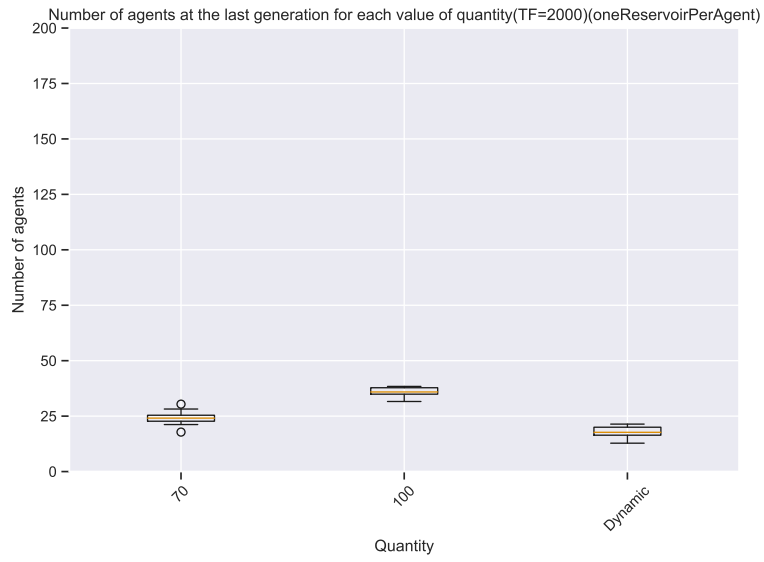


Figure 6: Nombre d'agents vivants sur la dernière génération en fonction de la quantité prise et lorsque l'environnement possède un TF de 2000

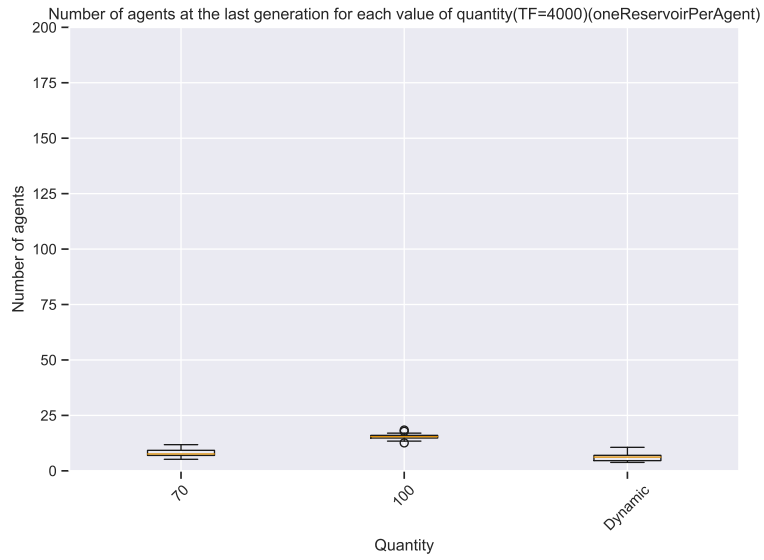


Figure 7: Nombre d'agents vivants à la dernière génération en fonction de la quantité prise et lorsque l'environnement possède un TF de 4000

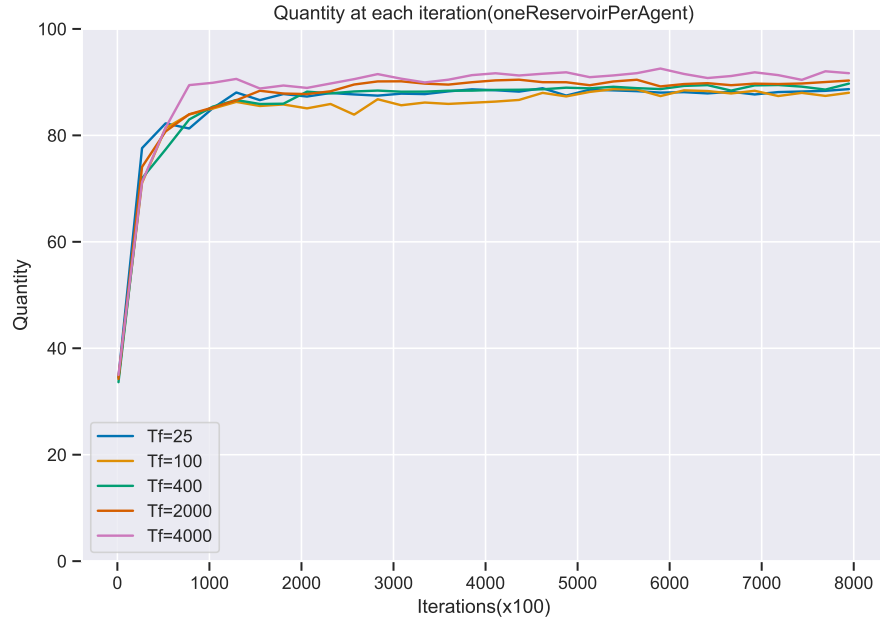


Figure 8: Quantité prise par les agents en fonction de la difficulté de l'environnement et du nombre d'itération dans l'expérience de Base

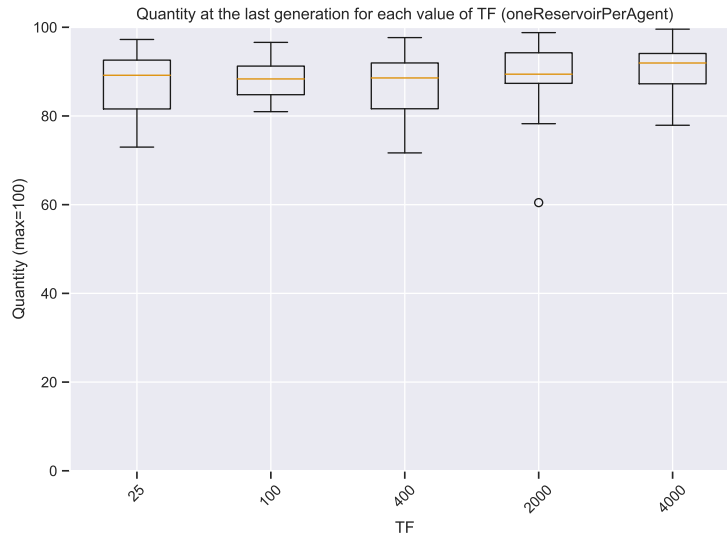


Figure 9: Boxplot représentant la quantité prise par les agents en fonction de la difficulté de l'environnement dans l'expérience de Base pour les dernières générations

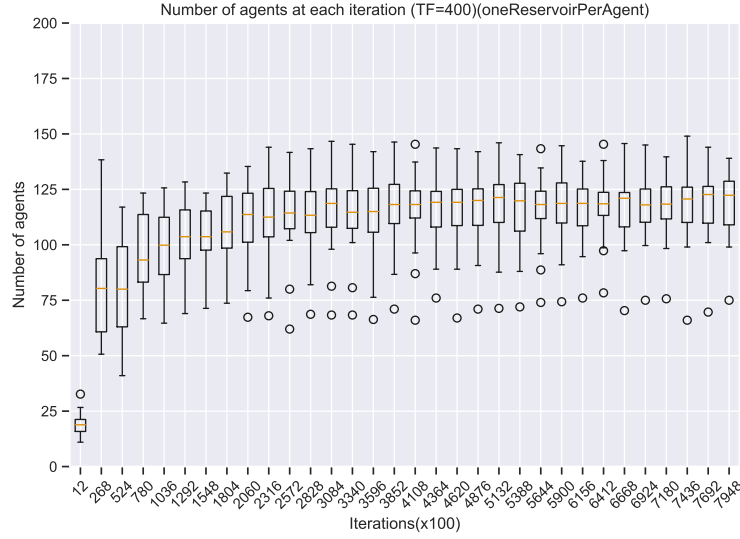


Figure 10: Nombre d'agents vivants à chaque génération et lorsque l'environnement possède un TF de 400, pour une quantité dynamique lors de l'expérience de base

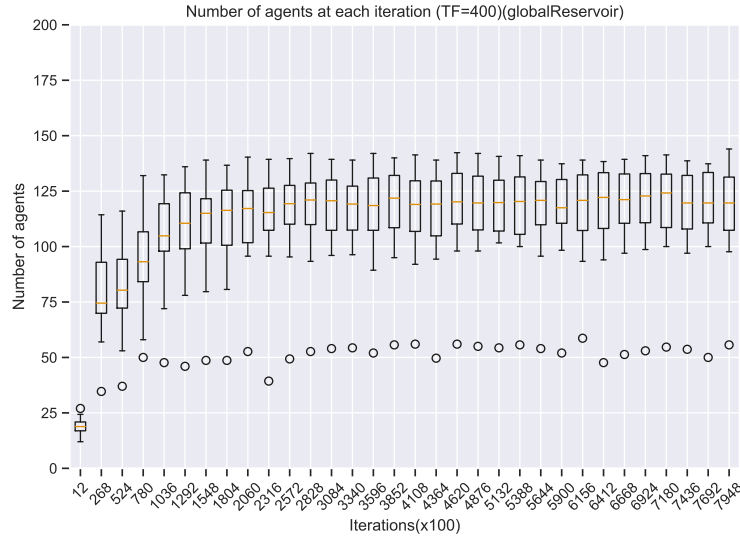


Figure 11: Nombre d'agents vivants à chaque génération et lorsque l'environnement possède un TF de 400, pour une quantité dynamique lors de l'expérience contrôle 1

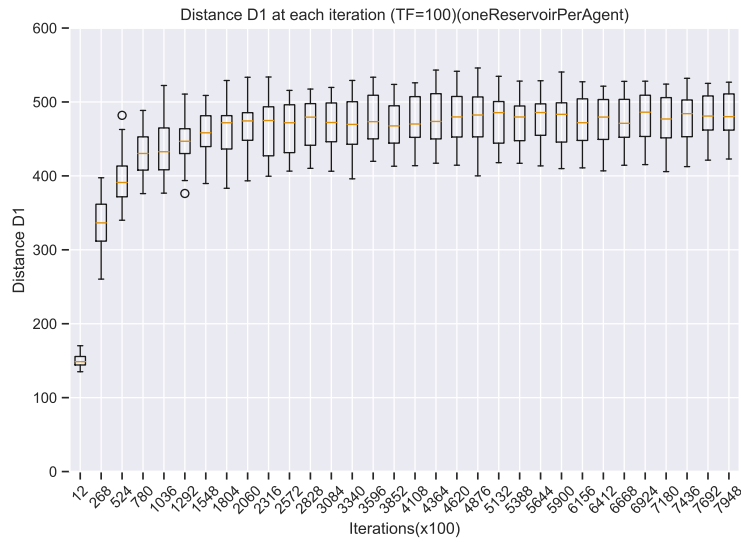


Figure 12: Distance D1 à chaque génération et lorsque l'environnement possède un TF de 100, pour une quantité dynamique lors de l'expérience de base

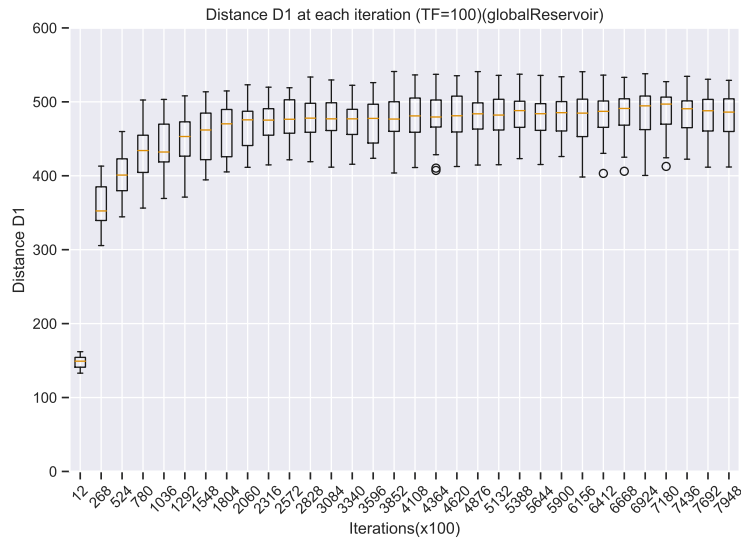


Figure 13: Distance D1 à chaque génération et lorsque l'environnement possède un TF de 100, pour une quantité dynamique lors de l'expérience contrôle 1

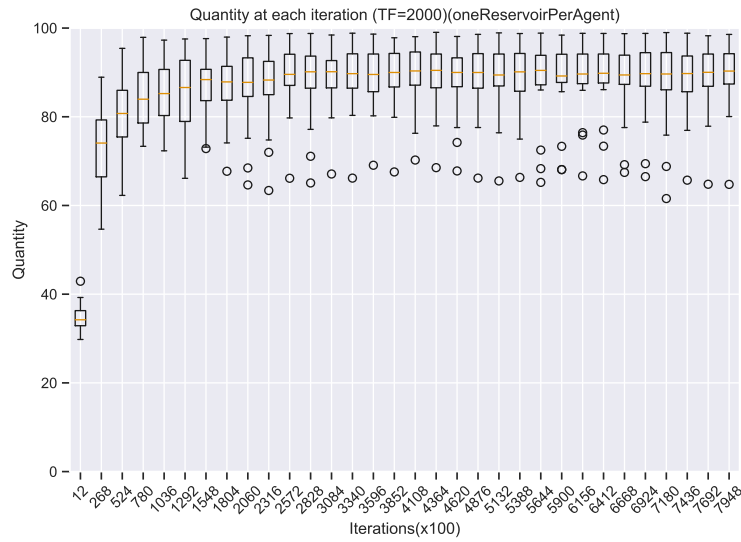


Figure 14: Quantité à chaque génération et lorsque l'environnement possède un TF de 2000, pour une quantité dynamique lors de l'expérience de base

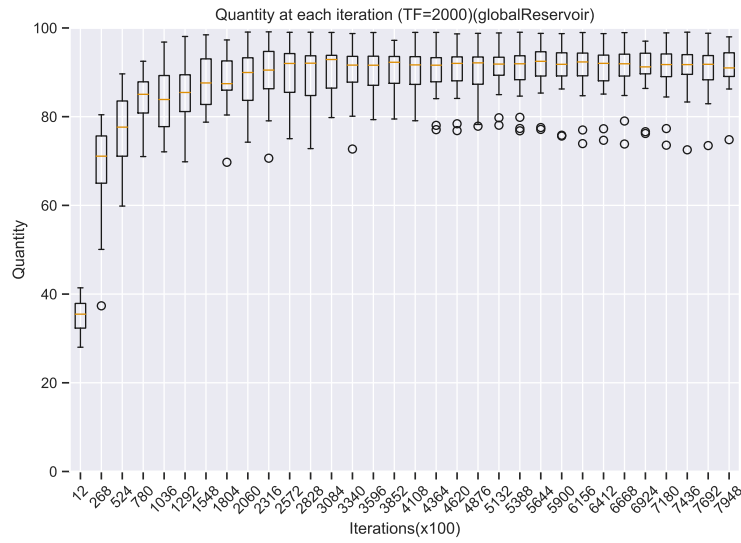


Figure 15: Quantité à chaque génération et lorsque l'environnement possède un TF de 2000, pour une quantité dynamique lors de l'expérience contrôle 1

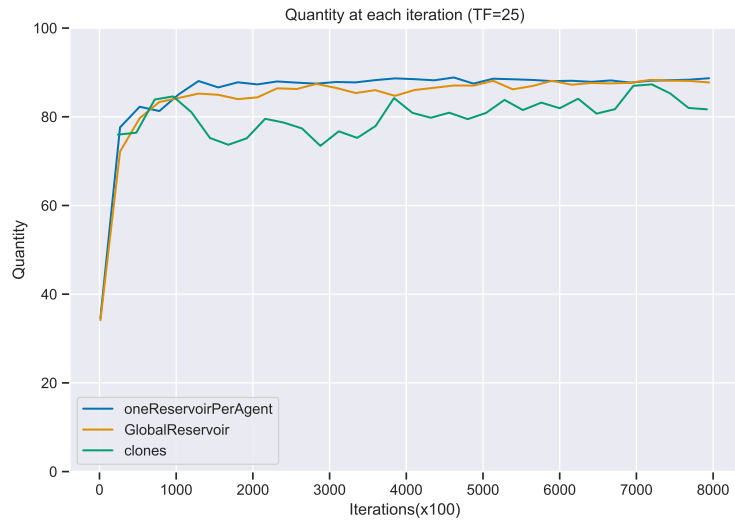


Figure 16: Comparaison des différentes quantités choisies lorsque l'environnement possède un TF de 25

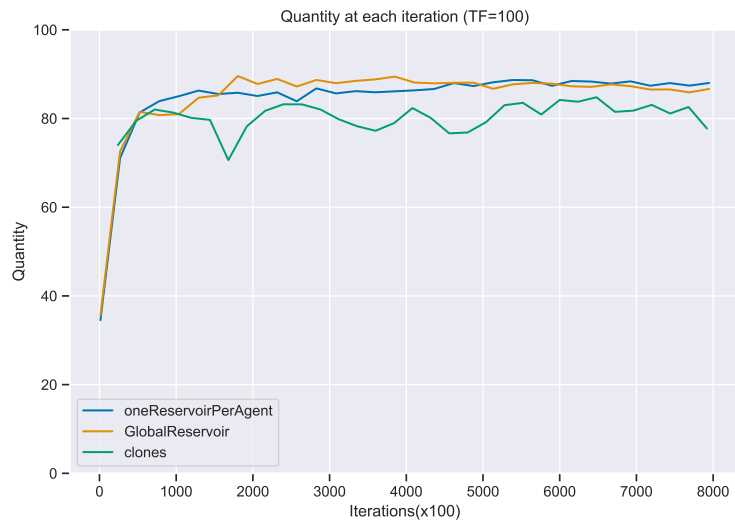


Figure 17: Comparaison des différentes quantités choisies lorsque l'environnement possède un TF de 100

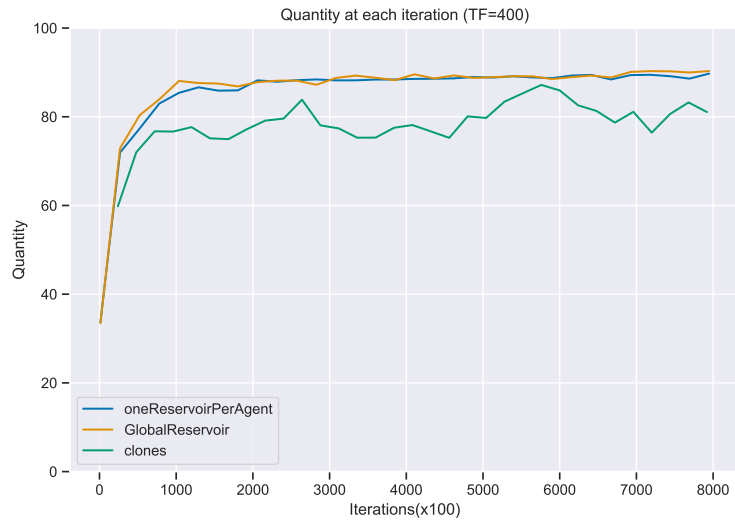


Figure 18: Comparaison des différentes quantités choisies lorsque l'environnement possède un TF de 400

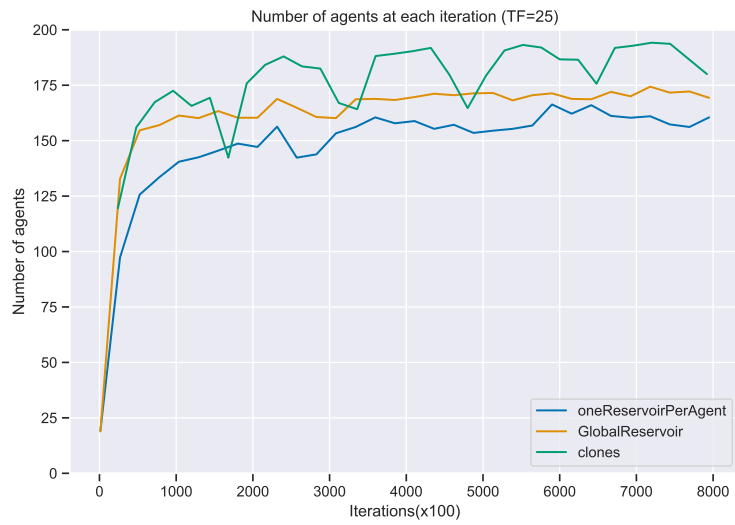


Figure 19: Comparaison du nombre d'agents vivants lorsque l'environnement possède un TF de 25

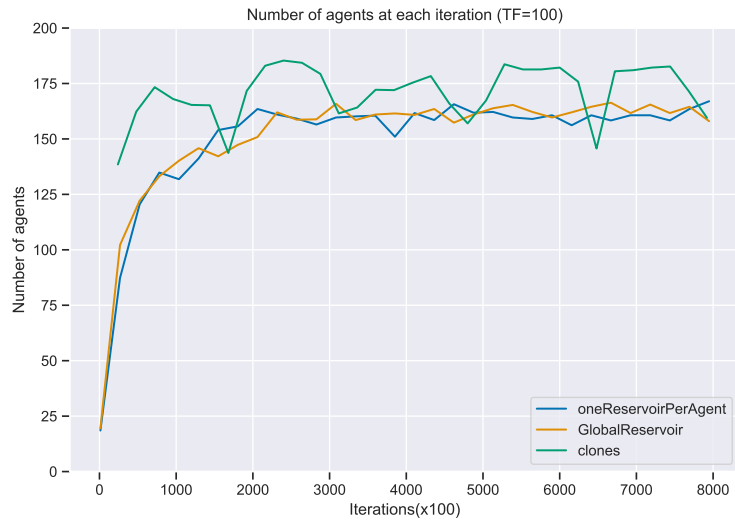


Figure 20: Comparaison du nombre d'agents vivants lorsque l'environnement possède un TF de 100

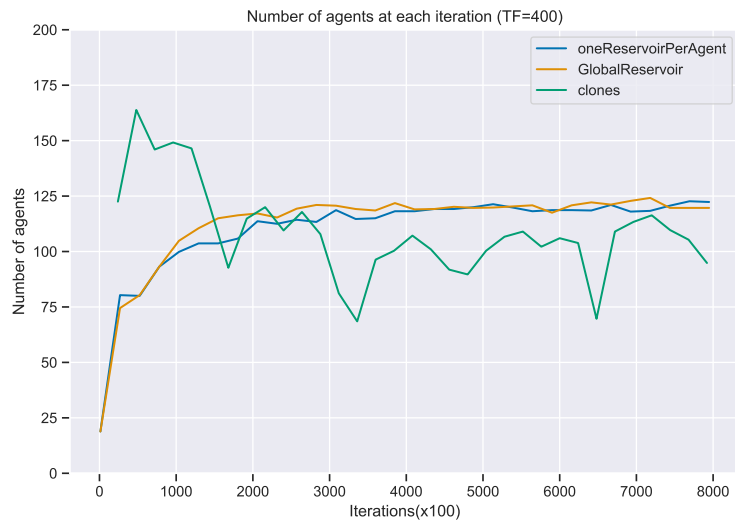


Figure 21: Comparaison du nombre d'agents vivants lorsque l'environnement possède un TF de 400

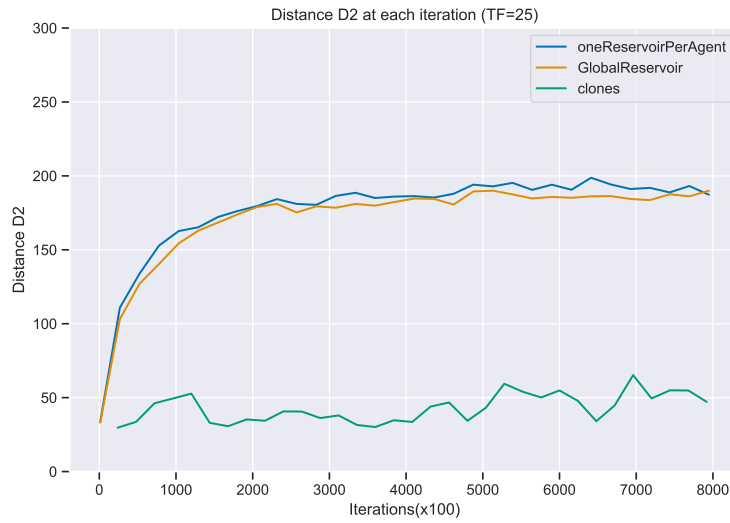


Figure 22: Comparaison des différentes valeurs de distance D2 lorsque l'environnement possède un TF de 25

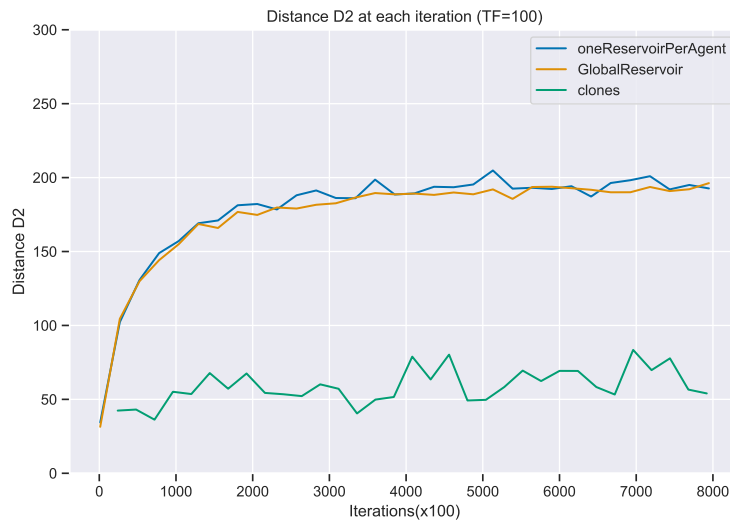


Figure 23: Comparaison des différentes valeurs de distance D2 lorsque l'environnement possède un TF de 100

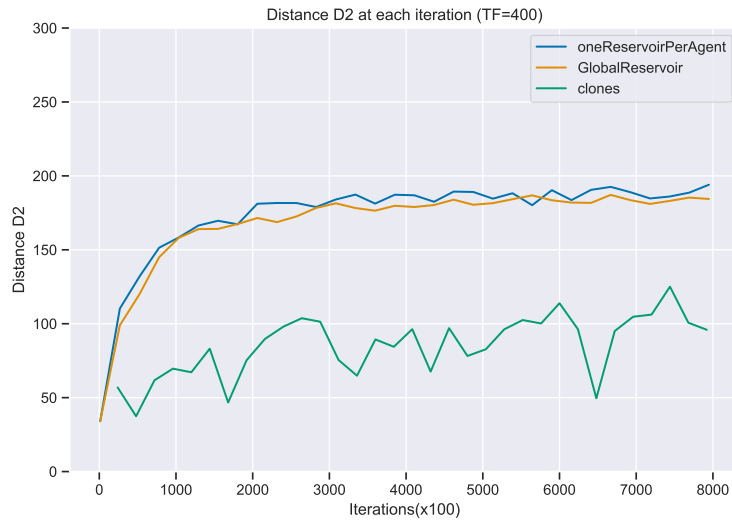


Figure 24: Comparaison des différentes valeurs de distance D2 lorsque l'environnement possède un TF de 400

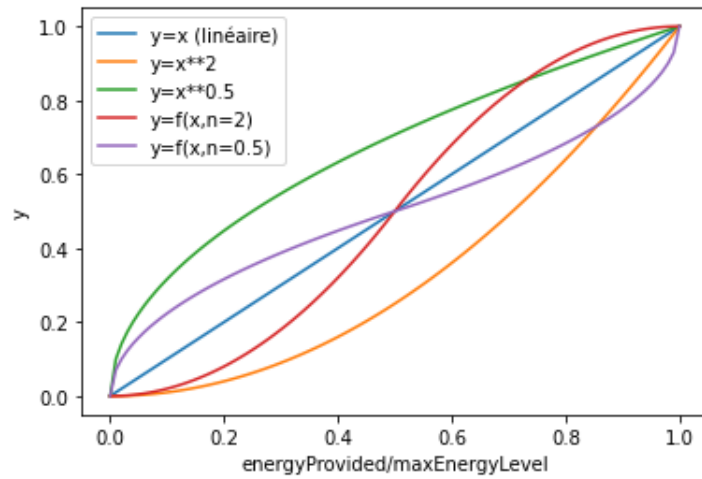


Figure 25: Fonctions possibles pour déterminer le temps de réapparition de la nourriture