

Université de Montréal

**Sur l'utilisation du langage de programmation Scheme pour le
développement de jeux vidéo**

par
David St-Hilaire

Département d'informatique et de recherche opérationnelle
Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures et postdoctorales
en vue de l'obtention du grade de Maître ès sciences (M.Sc.)
en informatique

Octobre, 2009

© David St-Hilaire, 2009.

Université de Montréal
Faculté des études supérieures et postdoctorales

Ce mémoire intitulé:

**Sur l'utilisation du langage de programmation Scheme pour le
développement de jeux vidéo**

présenté par:

David St-Hilaire

a été évalué par un jury composé des personnes suivantes:

Stefan Monnier
président-rapporteur

Marc Feeley
directeur de recherche

Pierre Poulin
membre du jury

Mémoire accepté le

RÉSUMÉ

Ce mémoire vise à recenser les avantages et les inconvénients de l'utilisation du langage de programmation fonctionnel dynamique Scheme pour le développement de jeux vidéo. Pour ce faire, la méthode utilisée est d'abord basée sur une approche plus théorique. En effet, une étude des besoins au niveau de la programmation exprimés par ce type de développement, ainsi qu'une description détaillant les fonctionnalités du langage Scheme pertinentes au développement de jeux vidéo sont données afin de bien mettre en contexte le sujet. Par la suite, une approche pratique est utilisée en effectuant le développement de deux jeux vidéo de complexités croissantes : *Space Invaders* et *Lode Runner*. Le développement de ces jeux vidéo a mené à l'extension du langage Scheme par plusieurs langages spécifiques au domaine et bibliothèques, dont notamment un système de programmation orienté objets et un système de coroutines. L'expérience acquise par le développement de ces jeux est finalement comparée à celle d'autres développeurs de jeux vidéo de l'industrie qui ont utilisé Scheme pour la création de titres commerciaux. En résumé, l'utilisation de ce langage a permis d'atteindre un haut niveau d'abstraction favorisant la modularité des jeux développés sans affecter les performances de ces derniers.

Mots clés : Langages de programmation fonctionnels, Scheme, jeux vidéo, programmation orientée objets, coroutines.

ABSTRACT

This master's thesis aims at pinpointing the pros and cons of using the dynamic functional language Scheme for developing video games. The method used is first based on a theoretical approach. Indeed, the specific requirements for video game programming and a detailed description of relevant Scheme features are presented. Then, a practical approach is taken by presenting two video games developed using the Scheme language: *Space Invaders* and *Lode Runner*. Their development resulted in the creation of various domain-specific languages and libraries, such as an object-oriented system and a coroutine system. Each of these are presented separately in their respective chapter. Finally, the experience achieved in this process is compared to the experience acquired by some video game companies that also used Scheme for the development of their titles. The use of Scheme allowed us to perform various high-level abstractions that improved the modularity of the video games developed, without affecting their performance.

Keywords: Functional programming languages, Scheme, video games, object-oriented programming, coroutines.

TABLE DES MATIÈRES

RÉSUMÉ	v
ABSTRACT	vii
TABLE DES MATIÈRES	ix
LISTE DES TABLEAUX	xv
LISTE DES FIGURES	xvii
REMERCIEMENTS	xxiii
CHAPITRE 1 : INTRODUCTION	1
1.1 Problématique	2
1.2 Méthodologie	3
1.3 Aperçu du mémoire	3
CHAPITRE 2 : DÉVELOPPEMENT DE JEUX VIDÉO	5
2.1 Historique	5
2.1.1 Préhistoire 1948-1970	5
2.1.2 Systèmes arcade 1970-1985	6
2.1.3 Premières consoles 1972-1984	6
2.1.4 Ordinateurs personnels 1977-...	7
2.1.5 Consoles portables 1980-...	8
2.1.6 Consoles intermédiaires 1984-2006	8
2.1.7 Consoles modernes 2005-...	9
2.2 Industrie du jeu vidéo	10
2.3 Contraintes de programmation	11
2.3.1 Fluidité	12
2.3.2 Modularité	13

2.3.3	Malléabilité	14
2.4	Conclusion	15
CHAPITRE 3 : LE LANGAGE SCHEME		17
3.1	Héritage Lisp	18
3.1.1	Histoire de Lisp	18
3.1.2	Avènement de Scheme	19
3.1.3	Langages inspirés de Scheme	20
3.2	Introduction au langage	20
3.2.1	Syntaxe	20
3.2.2	Aperçu des fonctionnalités requises par le standard	21
3.3	Programmation fonctionnelle	22
3.3.1	Programmation fonctionnelle pure	23
3.3.2	Programmation fonctionnelle en Scheme	24
3.4	Macros	26
3.5	Continuations	31
3.5.1	Exemples d'utilisations	33
3.5.2	Conclusion	36
3.6	Gestion mémoire automatique	36
3.6.1	Survol des techniques	37
3.6.2	Conclusion	40
3.7	Dynamisme du langage	40
3.7.1	Interprétation et débogage efficace	41
3.8	Extensions utiles	42
3.8.1	Fermetures	42
3.8.2	Tables de hachage	43
3.8.3	Tests d'égalité et opérations conditionnelles	44
3.8.4	Déclaration et utilisation d'objets	45
3.8.5	Affichage de résultats	45

3.8.6	Interfaces aux fonctions étrangères	46
3.8.7	Capture de continuations	47
3.9	Conclusion	47
CHAPITRE 4 : PROGRAMMATION ORIENTÉE OBJET		49
4.1	Description du système	53
4.1.1	Définition de classes	54
4.1.2	Définition de fonctions génériques	57
4.1.3	Fonctions et formes spéciales utilitaires	62
4.2	Implantation	64
4.2.1	Implantation de <code>define-class</code>	65
4.2.2	Implantation de <code>define-generic</code>	69
4.2.3	Implantation de <code>define-method</code>	71
4.3	Performances	72
4.4	Conclusion	74
CHAPITRE 5 : SYSTÈME DE COROUTINES		77
5.1	Description du langage	79
5.1.1	Création de coroutines	80
5.1.2	Démarrage du système	81
5.1.3	Manipulation du flot de contrôle	84
5.1.4	Environnement dynamique	86
5.1.5	Système de communication inter-coroutines	86
5.1.6	Autres mécanismes de synchronisation	90
5.2	Implantation	90
5.2.1	Implantation des coroutines	91
5.2.2	<i>Timers</i>	92
5.2.3	Ordonnancement	93
5.2.4	Système de messagerie	97

5.3	Performances	100
5.4	Conclusion	103
CHAPITRE 6 : ÉVALUATION ET EXPÉRIENCES		105
6.1	Développement de <i>Space Invaders</i>	106
6.1.1	Version initiale	107
6.1.2	Version orientée objets	116
6.1.3	Version avec flot de contrôle sous forme de coroutines	119
6.1.4	Conclusion	122
6.2	Développement de <i>Lode Runner</i>	123
6.2.1	Logique du jeu	124
6.2.2	Gestion de l'état des objets	125
6.2.3	Détection et résolution de collisions	130
6.2.4	Rendu graphique	131
6.2.5	Intelligence artificielle	132
6.2.6	Conclusion	134
6.3	Performances	134
6.4	Conclusion	136
CHAPITRE 7 : TRAVAUX RELIÉS		139
7.1	Comparaison de langages	139
7.1.1	C++	139
7.1.2	Lua	141
7.2	Sommaire	142
7.3	Jeux en Lisp	144
7.3.1	Naughty Dog	144
7.3.2	<i>QuantZ</i>	146
7.4	Autres version des jeux développés	147
7.4.1	<i>Space Invaders</i>	148

7.5 Conclusion	150
CHAPITRE 8 : CONCLUSION	151
BIBLIOGRAPHIE	155

LISTE DES TABLEAUX

4.1	Comparaison de la présence des fonctionnalités désirées par les implan- tations existantes de systèmes d'objets disponibles sur Gambit-C	53
4.2	Complexité algorithmique des opérations du système objet effectuées du- rant l'exécution	72
4.3	Temps d'exécution (en secondes) de 1 000 000 itérations de tests com- paratifs sur les opérations du système de programmation orientée objets développé avec d'autres systèmes objet existants	73
4.4	Comparaison de performances relatives au système le plus rapide (1.0) pour chaque opération du système objet	74
5.1	Temps d'exécution (en secondes) de 1 000 000 itérations de tests compa- ratifs entre le système développé et d'autres systèmes similaires	101
5.2	Comparaison de performances relatives au système le plus rapide (1.0) en fonction des données du tableau 5.1	102
6.1	Analyse de performances des jeux développés	135
7.1	Comparaison des langages C++, Lua et Scheme	143
7.2	Comparaison entre les versions originales et réécrites des jeux développés.	149

LISTE DES FIGURES

3.1	Exemple simple d'appels de fonctions en Scheme	21
3.2	Exemples d'utilisation de listes sous forme de S-expression	21
3.3	Exemples d'utilisation des types de base de Scheme	22
3.4	Définition de la fonction <code>flip</code> en lambda calcul non typé	23
3.5	Implantation de paires en utilisant des fermetures du lambda calcul . .	24
3.6	Exemple de création et d'utilisation de fonctions d'ordre supérieur. . . .	25
3.7	Fonction d'ordre supérieur <code>foldl</code>	25
3.8	Utilisation d'un <i>thunk</i> dans le but de retarder l'évaluation de code . . .	26
3.9	Exemple de macro textuelle en C	26
3.10	Exemple simple de macro utilisant le passage par nom à profit	28
3.11	Exemple de macro ajoutant une nouvelle forme spéciale au langage . .	28
3.12	Résultat de l'expansion macro de l'appel présenté dans la figure 3.11 . .	28
3.13	Macro <code>for</code> hygiénique	29
3.14	Exemple plus avancé de macro Scheme	30
3.15	Macro venant augmenter le petit <i>DSL</i> portant sur la manipulation de masques binaires	30
3.16	Implantation simple de la fonction factorielle	31
3.17	Forme CPS de la fonction factorielle	32
3.18	Réification d'une continuation à l'aide <code>call/cc</code>	33
3.19	Utilisation non efficace de <code>foldl</code>	34
3.20	Utilisation d'une continuation pour échapper au flot de contrôle	35
3.21	Exemple de recherche par retour arrière	35
3.22	Utilisation du passage de paramètres par mots clés	43
3.23	Exemple d'utilisation du paramètre <i>reste</i>	43
3.24	Exemples d'utilisations de la fonction <code>apply</code>	43

3.25	Exemple d'utilisation d'une table de hachage spécialisée pour des clés symboliques et d'utilisation de la forme spéciale cond	44
3.26	Exemple de création d'objets dans Gambit-C	45
3.27	Exemples d'utilisations de fonctions d'affichage	46
3.28	Exemple d'utilisation de l'interface à des fonctions étrangères de Gambit-C	46
3.29	Exemple d'utilisation de l'API améliorée de capture de continuations fournie par Gambit-C	47
4.1	Exemple d'utilisation de méthodes à la manière traditionnelle (en Java)	51
4.2	Exemple de <i>dispatch multiple</i> écrit en CLOS	52
4.3	Syntaxe de la forme spéciale define-class	54
4.4	Utilisation simple du système objet compatible avec la forme define-type	55
4.5	Exemple d'utilisation dynamique d'attributs de classe	56
4.6	Utilisation générique de constructeurs	57
4.7	Exemple de déclarations et d'utilisations d'une fonction générique . . .	58
4.8	Exemple de discrimination d'instances de fonctions génériques par la valeur des arguments	60
4.9	Exemple illustrant l'algorithme de sélection d'instances de fonctions génériques	61
4.10	Exemple d'utilisation de <i>cast</i> pour l'appel de fonction générique	62
4.11	Exemple d'utilisation des formes spéciales fournies par le système objet	64
4.12	Exemple concret de structure d'instance, descripteur de classe et de fonction d'accès à un attribut par indirection	66
4.13	Schématisation des structures créées dans la figure 4.12	66
4.14	Code généré pour l'accès à l'attribut x de la classe point de la figure 4.12	68
4.15	Code expansé effectuant le <i>dispatch</i> dynamique pour la fonction générique init!	70
4.16	Implantation de la recherche d'instances polymorphiques d'une fonction générique	71
5.1	Exemple générique d'utilisation de systèmes de coroutines cascades . .	78

5.2	Architecture globale du système de coroutines	80
5.3	Diagramme d'états d'une coroutine	81
5.4	Exemple de démarrage d'un système de coroutines	82
5.5	Exemple de démarrage de système de coroutines en cascade	84
5.6	Exemple de modification de la continuation d'une coroutine	86
5.7	Exemple de patrons pouvant être utilisés dans le filtrage par motifs de la forme spéciale recv	88
5.8	Structure de données représentant une coroutine	91
5.9	Obtention de la valeur de retour d'une coroutine	92
5.10	Algorithme d'ordonnancement	95
5.11	Fonctions de changement de contexte explicite	96
5.12	Implantation du retour au contexte d'une coroutine	96
5.13	Envoi de messages entre coroutines	97
5.14	Réception de messages avec la fonction ?	98
5.15	Illustration des priorités de réception de messages avec la forme spéciale recv	99
5.16	Expansion macro du corps de la coroutine c1 de la figure 5.15	100
5.17	Patron de filtrage utilisé pour les tests de performance. Le message reçu par ce dernier est celui attendu par la deuxième branche	101
5.18	Test de performance ping-server utilisé pour comparer une utilisation réaliste de la réception de messages avec filtrage de motifs	101
6.1	Capture d'écran du jeu <i>Space Invaders</i> développé	106
6.2	Exemple de fonte utilisée dans <i>Space Invaders</i>	109
6.3	Exemple de code de simulation par événements discrets	110
6.4	Évènement représentant le vaisseau ennemi de type <i>mothership</i>	111
6.5	Exemple d'animation combinant les événements discrets et une program- mation CPS	112
6.6	Structures de données utilisées dans la première version de <i>Space Invaders</i>	113

6.7	Détection de collision dans <i>Space Invaders</i>	114
6.8	Résolution de collisions dans la première version de <i>Space Invaders</i> . .	114
6.9	Évènement de mort d'un joueur	115
6.10	Aperçu de la hiérarchie de classe du jeu	117
6.11	Exemple de déclarations d'instances de la fonction générique resolve-collision! dans <i>Space Invaders</i>	119
6.12	Classe de base des objets en tant que coroutines	120
6.13	Comportement du méta-objet invader-controller qui régit le compor- tement d'une rangée d' <i>invaders</i>	122
6.14	Capture d'écran du jeu <i>Lode Runner</i> développé	124
6.15	Boucle principale du jeu <i>Lode Runner</i>	125
6.16	Définition de la classe human-like dans le jeu <i>Lode Runner</i>	126
6.17	Gestion de l'état d'une entité appartenant à la hiérarchie de type human-like dans <i>Lode Runner</i>	127
6.18	Forme spéciale utilisée afin de simplifier la définition d'animations cy- cliques d'entités du jeu <i>Lode Runner</i>	127
6.19	Définition d'animation cyclique dans le jeu <i>Lode Runner</i>	128
6.20	Diagramme d'états d'un niveau de <i>Lode Runner</i>	128
6.21	Interface du système de machines à états développé pour les niveaux de <i>Lode Runner</i>	129
6.22	Implantation de la machine à états pour les niveaux de <i>Lode Runner</i> .	129
6.23	Action de transition vers l'état start d'un niveau du jeu <i>Lode Runner</i>	130
6.24	Détection de collisions dans <i>Lode Runner</i>	131
6.25	Implantation des couches de rendu dans <i>Lode Runner</i>	132
6.26	Algorithme trivial d'intelligence artificielle des robots	132
6.27	Algorithme d'intelligence artificielle reproduisant un comportement <i>si-</i> <i>miltaire</i> au jeu original	133
6.28	Utilisation dynamique des algorithmes d'intelligence artificielle	133

7.1	Exemple de code en GOAL utilisant le système de coroutines	145
7.2	Déclaration de données en Scheme utilisées dans le jeu <i>Uncharted : Drake's fortune</i> [1]	146

REMERCIEMENTS

Je tiens à remercier ma femme Julie, mon fils Émile ainsi que ma famille proche pour le soutien précieux qu'ils m'ont apporté tout au long de mes études universitaires. Je n'aurais jamais pu y arriver sans leur aide.

Je voudrais également remercier tous les membres du laboratoire de traitement parallèle dans lequel j'ai effectué les travaux de recherche présentés dans ce mémoire. Leurs connaissances et amitié me suivront tout au long de mon cheminement professionnel. Mon directeur de recherche, Marc Feeley, a joué un rôle très important dans ces études, tant au niveau des connaissances scientifiques qu'au niveau de la persévérance et de la continuité. Je lui en suis très reconnaissant.

Je tiens également à remercier le Conseil de recherche en sciences et en génie du Canada (CRSNG) pour son précieux soutien financier qui m'a permis d'entreprendre la recherche présentée dans ce document.

Aussi, je voudrais remercier le studio de développement de jeux vidéo *Gamerizon Inc.* pour lequel je travaille et qui m'a permis de prendre congé afin de rédiger ce mémoire de maîtrise.

Enfin, je voudrais remercier tous les lecteurs et lectrices qui auront le courage de lire ce document au complet !

CHAPITRE 1

INTRODUCTION

L'industrie du jeu vidéo devient de plus en plus importante dans le domaine de l'informatique. Cette croissance est bien reflétée par l'augmentation de 28 % des revenus provenant de la vente de jeux vidéo aux États-Unis durant l'année 2007 [2]. La place occupée par l'industrie du jeu vidéo durant cette même année s'estime à 76 % du marché de tous les logiciels vendus [3].

L'engouement du marché du jeu vidéo incite les compagnies oeuvrant dans le domaine à repousser les limites de l'état de l'art du développement de jeux. La compétition est féroce et beaucoup d'efforts doivent donc être investis dans la création d'un jeu afin qu'il se démarque de la masse et devienne un nouveau *Blockbuster Hit*. Le président de la compagnie française UbiSoft estime que le coût moyen de développement d'un jeu sur une console moderne se situe entre 20 et 30 millions de dollars [4]. Si l'on estime le salaire moyen d'un artiste ou d'un dév, cela reviendrait à un travail d'environ 300 à 500 années*hommes.

Puisque la création d'un jeu vidéo peut nécessiter autant d'efforts, il semble très intéressant de faciliter le développement de ces derniers. Avec autant d'efforts mis en place, même une petite amélioration sur le cycle de développement peut engendrer une diminution énorme des coûts de production et améliorer la qualité des environnements utilisés par les développeurs.

Jusqu'à ce jour, la grande majorité des jeux vidéo sont écrits à l'aide de langages relativement de bas niveau, tels que C [5], C++ [6] ou encore C# [7]. Les langages de plus bas niveaux sont caractérisés par un modèle d'implantation près de l'architecture de la plate-forme sur laquelle le logiciel est déployé. Ces langages sont généralement utilisés parce qu'ils sont déjà bien établis et que la main d'oeuvre est facilement accessible.

Les langages de haut niveau sont, quant à eux, caractérisés par une bonne abstraction du système utilisé. Ces derniers permettent d'exprimer un calcul de manière

naturelle. Ils facilitent donc le travail de programmation. Le coût de ces abstractions se répercute généralement en un coût de performance du programme. Dans le passé, la performance était critique dans les jeux vidéo, mais les consoles modernes sont devenues plus performantes que la plupart des ordinateurs personnels. Ainsi, la performance n'est donc plus aussi critique qu'auparavant. De plus, les améliorations du domaine de la compilation de langages de haut niveau font en sorte que les performances de ces systèmes sont comparables à l'utilisation de langages de plus bas niveau.

Ainsi, l'utilisation de langages de haut niveau pourrait potentiellement réduire le temps de développement des jeux en permettant aux programmeurs et designers de s'exprimer plus facilement. Les langages de la famille Lisp semblent être de bons candidats en tant que langages de haut niveau. Le langage Scheme [8], un Lisp moderne et performant, offre ainsi plusieurs fonctionnalités de haut niveau. On retrouve, entre autres, du typage dynamique, des fonctions de première classe, un système de macros évolué, un accès direct aux continuations du calcul et un système de débogage dynamique. Ces particularités du langage Scheme sont discutées plus en détail dans le chapitre 3.

Le système Gambit-C [9], l'une des implantations de Scheme les plus performantes [10], sera utilisé pour effectuer les expériences pratiques. Ce système comporte de nombreuses extensions pouvant être très utiles au développement de jeux vidéo. On y retrouve entre autres des tables de hachage et des *threads* (processus légers).

Une utilisation judicieuse de ce système pourrait potentiellement bénéficier à des projets aussi complexes que le sont les productions de jeux vidéo.

1.1 Problématique

Ce mémoire de maîtrise vise à répondre à la problématique suivante :

Quelles sont les forces et les faiblesses du langage de programmation Scheme pour le développement de jeux vidéo ?

1.2 Méthodologie

Afin de répondre à la problématique posée, nous avons étudié les caractéristiques de Scheme et du compilateur Gambit-C, ainsi que les besoins au niveau du développement de jeux vidéo. Par la suite, deux jeux ont été développés en exploitant ces caractéristiques dans le but de raffiner nos approches et les évaluer dans un contexte pratique.

Le premier jeu a servi de plate-forme d'exploration permettant d'élaborer une méthodologie efficace pour le développement de jeux. Afin d'obtenir une telle méthodologie, plusieurs itérations de développement ont été effectuées, chacune permettant d'explorer de nouveaux aspects sur la manière de résoudre des problématiques associées à la création de jeux. Par exemple, la synchronisation des entités dans le jeu ou la description efficace d'un système de détection et de résolution de collisions ont été résolues.

Suite à l'écriture de ce jeu, un deuxième jeu plus complexe a été développé. Ce dernier visait à consolider les techniques précédemment utilisées et étendre celles-ci dans le cadre de ce jeu qui présente de nouveaux défis, tel que l'implantation d'une intelligence artificielle.

1.3 Aperçu du mémoire

Ce mémoire est composé de quatre parties. La première partie est une présentation de l'industrie des jeux vidéo et des défis découlant du développement de ces derniers. Par la suite est présentée une introduction traitant de programmation fonctionnelle, du langage de programmation Scheme et des outils de développement disponibles, en particulier le compilateur Gambit-C. Ce chapitre donne un aperçu général de ces langages et permet de saisir les concepts fondamentaux du langage Scheme.

La deuxième partie du mémoire porte sur des extensions faites au langage Scheme qui ont été utilisées dans le but d'améliorer le développement de jeux vidéo. On y présente la programmation orientée objet et un système d'objets conçu pour répondre aux besoins de la programmation de jeux vidéo. Un système de coroutines conçu dans

les mêmes optiques est également présenté.

La troisième partie porte sur l'expérience acquise par l'auteur en effectuant l'écriture de deux jeux vidéo simples, mais possédant suffisamment de complexité pour exposer les problèmes associés au développement de jeux vidéo en général. De plus, la façon de tirer profit du langage de programmation Scheme pour résoudre ces problèmes est abordée. Une présentation des travaux reliés aux résultats présentés suit ce dernier chapitre. On y parle de l'utilisation d'autres langages pour le développement de jeux vidéo. On y cite aussi des exemples d'utilisation des langages Scheme et Lisp dans des jeux vidéo commerciaux et de l'expérience tirée de cette utilisation par les développeurs.

Finalement, une conclusion apporte la lumière sur la problématique exposée dans ce mémoire. Le point sur l'expérience acquise pour le développement de jeux vidéo en Scheme y est fait et les avantages et inconvénients ou problèmes rencontrés sont exposés.

CHAPITRE 2

DÉVELOPPEMENT DE JEUX VIDÉO

Les jeux vidéo font partie d'un domaine de l'informatique en pleine effervescence grâce à une demande constante de nouveaux produits.

L'histoire des jeux vidéo possède déjà un demi-siècle de créations de tous genres qui ont contribué à générer l'engouement actuel pour ces derniers. Un bref historique des jeux vidéo est présenté dans ce chapitre afin d'illustrer l'évolution de ceux-ci et de situer dans le temps les jeux développés pour ce mémoire.

L'industrie du jeu vidéo actuelle est le moteur qui permet aux jeux vidéo de continuer à évoluer et se raffiner. Un aperçu global de cette industrie est ainsi donné dans ce chapitre.

Finalement, les besoins au niveau de la programmation de jeux vidéo sont énoncés afin de cerner les composantes importantes pour un langage de programmation utilisé pour le développement de jeux.

2.1 Historique

L'histoire des jeux vidéo s'étale sur environ un demi-siècle. Cette dernière est une source importante d'informations qui permet d'exposer l'évolution des jeux vidéo. Seulement un bref résumé est présenté afin de donner les grandes lignes de l'évolution des jeux au cours des cinquante dernières années [11, 12].

2.1.1 Préhistoire 1948-1970

Les jeux vidéo ont fait leur apparition avant même les premiers ordinateurs. En effet, le *Cathode Ray Amusement Device* [13] fit son apparition en 1948. Il s'agissait d'un jeu rendu sur un tube cathodique simulant le lancement de missiles sur des cibles.

Vers le début des années 1960, quelques jeux ont fait leur apparition sur les premiers ordinateurs universitaires, notamment au *MIT* (Massachusetts Institute of Technology)

et à l'université de Cambridge. On y retrouve par exemple le jeu *Spacewar!* qui est l'un des premiers jeux multi-joueurs mettant les joueurs en adversité dans leurs vaisseaux spatiaux pouvant lancer des missiles. Ce jeu a d'ailleurs été une motivation pour les premiers efforts de développement du système d'exploitation UNICS (UNiplexed Information and Computing Service) [14].

Avec un intérêt soutenu envers le potentiel de divertissement qu'offrait les jeux vidéo de l'époque, le développement de machines dédiées aux jeux vidéo a ainsi débuté.

2.1.2 Systèmes arcade 1970-1985

En 1971, des étudiants de l'université de Stanford ont réimplanté le jeu *Spacewar!* sur une machine fonctionnant avec de la monnaie. Ce fut la première machine arcade.

Par la suite, le développement de telles machines est devenu très répandu. En 1972, la compagnie *Atari* fut fondée et démarra l'industrie du jeu vidéo sur arcade avec le jeu *Pong*. Il s'agissait d'un jeu de tennis de table permettant à deux joueurs de se mesurer l'un à l'autre. Ce jeu demeure toujours très célèbre.

Ce fut alors l'âge d'or des machines arcade où l'on a créé beaucoup de jeux aux contenus diversifiés. Parmi ceux-ci, *Space Invaders* (1978) et *PacMan* (1980) furent extrêmement populaires.

Malgré la simplicité des jeux arcade, ils étaient beaucoup appréciés par les joueurs qui tentaient continuellement de se surpasser.

De nos jours, on retrouve toujours des salles arcades, mais celles-ci ont perdu de la popularité puisque les consoles de jeux se retrouvent dans nos salons.

2.1.3 Premières consoles 1972-1984

La première console de jeu vidéo, le *Magnavox Odyssey* fit son apparition en Amérique du Nord en 1972. Cette console permettait de jouer sur la télévision, assis confortablement dans le salon. Elle permettait aussi d'insérer des jeux sous forme de cartouches. Un total de 28 jeux étaient disponibles pour cette console, qui malgré l'absence de

périphérique audio, a été vendue en 300 000 exemplaires.

Plusieurs autres consoles ont par la suite été créées, allant d’une version console de *Pong* jusqu’à des consoles plus puissantes comme le *ColecoVision* qui démarrèrent l’ère des consoles 8-bits. La vocation principale de ces dernières n’étaient pas d’innover dans le développement de jeux, mais plutôt de rendre disponible les jeux arcades populaires.

2.1.4 Ordinateurs personnels 1977-...

Les premiers ordinateurs personnels firent leur apparition vers la fin des années 1970, notamment l’ordinateur *Apple II* produit par *Apple Inc.* Ces ordinateurs personnels offraient plus de puissance que les consoles de l’époque et donnaient la possibilité aux amateurs de créer leurs propres jeux. Les jeux d’ordinateurs étaient distribués sur beaucoup de média différents. La distribution allait de cassettes, aux disquettes, en passant bien sûr par des échanges postaux de code sources. En 1982, le jeu *Lode Runner* fut développé pour les ordinateurs *Apple II*. Ce jeu d’action fut l’un des premiers comprenant un éditeur de niveaux.

Aussi, le jeu *Rogue* fut créé durant les années 1980, pour les premiers systèmes Unix. Il fut le pionnier d’un nouveau genre de jeux (surnommé *Roguelike*) qui différait beaucoup des jeux d’action retrouvés en salles d’arcade ou sur console. Il présentait une interface visuelle très minimaliste. La narration, qui était un point central du jeu, était effectuée de manière textuelle et le joueur interagissait de même avec le jeu.

Les jeux d’ordinateurs ont longtemps été supérieurs aux jeux de consoles puisque les ordinateurs étaient toujours à la fine pointe de la technologie, et les consoles accusaient un retard à ce niveau. Ainsi, on retrouvait des jeux ayant de meilleurs graphiques ou utilisant des périphériques plus variés sur les ordinateurs personnels. Par conséquent, les jeux d’ordinateurs existaient dans un monde parallèle à celui des consoles et ne se livraient pas de compétition réelle.

Par contre, avec l’avènement des consoles modernes qui sont plus performantes que la plupart des ordinateurs personnels, cet énoncé n’est plus valide. Il en résulte que les

jeux sur des consoles actuelles surpassent souvent les jeux d'ordinateurs.

2.1.5 Consoles portables 1980-...

Les toutes premières consoles portables furent développées par la compagnie *Mattel Toys* qui créèrent les jeux *Auto Race* et *Football*. Ces derniers étaient distribués sur des consoles de la taille d'une calculatrice, ne nécessitant pas de téléviseurs. Chaque console était dédiée à un seul jeu. Celles-ci furent un succès rapportant plusieurs centaines de millions de dollars à leurs créateurs.

Par la suite, les grandes compagnies du jeu vidéo comme *Nintendo* et *Bandai* se sont intéressées à de telles consoles et produisirent plusieurs exemplaires des consoles *Game & Watch* qui contenaient des succès d'arcade tels *Mario's Cement Factory* et *Donkey Kong Jr.*

Le même concepteur de *Game & Watch* a par la suite fusionné ces dernières avec le contrôleur du *Nintendo Entertainment System* (NES) pour obtenir la première console portable à grand succès : le *GameBoy*. Cette console qui offrait un écran monochrome fut vendue à 118 millions d'exemplaires. Le jeu le plus vendu sur cette console fut nulle autre que *Tetris* qui a été vendu à environ 33 millions d'unités.

Les consoles portables ont continué d'évoluer grandement et elles sont actuellement équivalentes aux consoles d'une ou deux générations précédentes. Par exemple, le Sony *PlayStation Portable* rivalise en matière de puissance de matériel à la précédente console de Sony, le *PlayStation 2*. On peut donc développer des jeux très complexes sur ces consoles portables, mais ils ne peuvent toujours pas rivaliser avec les jeux sur consoles ou ordinateurs.

2.1.6 Consoles intermédiaires 1984-2006

Au début des années 1980, plusieurs consoles étaient disponibles sur le marché, mais une saturation de mauvais jeux et des problèmes de gestion ont fait en sorte que l'industrie du jeu vidéo a connu une grande dépression. Cette dernière a pris subitement

fin avec la sortie de la console produite par *Nintendo*, le *Nintendo Entertainment System* (NES), qui connut un succès fulgurant en vendant 62 millions de consoles dans le monde. La grande qualité des jeux produits a certainement favorisé l'adoption de cette nouvelle console. On retrouve entre autres des jeux de tous genres comme le jeu de plate-forme *Super Mario Bros*, le jeu d'aventure *The Legend of Zelda* et le jeu d'action/aventure *Metroid*.

Par la suite, la compagnie SEGA sortit un compétiteur sérieux au NES, le *Master System*, qui rivalisait en termes de puissance matérielle et de prix. Depuis ce temps se livre la guerre des consoles où lorsqu'une compagnie développe un nouveau produit, les compétiteurs ne tardent pas à en produire un similaire.

2.1.7 Consoles modernes 2005-...

Les consoles modernes sont généralement conçues avec du matériel à la fine pointe de la technologie et sont adaptées à un certain public cible.

Au moment de l'écriture de ce mémoire, on retrouve la console *Wii* de Nintendo conçue pour un public constitué de joueurs occasionnels grâce au contrôleur innovateur de cette console permettant de jouer en effectuant des mouvements plus naturels.

D'un autre côté compétitionnent le *PlayStation 3* de Sony et le *XBox 360* de Microsoft qui cherchent à convaincre les joueurs plus assidus d'utiliser leur système en offrant des consoles à la fine pointe de la technologie et à prix très raisonnable en comparaison des prix des ordinateurs de jeux équivalents. Par exemple, le *PlayStation 3* utilise une architecture *Cell* qui comprend une unité principale de 3,2 GHz et 8 sous-processeurs moins puissants permettant d'exécuter plusieurs processus en parallèle. Ce dernier possède aussi une carte graphique puissante comprenant une mémoire vidéo de 256 Mo.

2.2 Industrie du jeu vidéo

Tel que mentionné dans le chapitre 1, l'industrie du jeu vidéo est très importante et génère des milliards de dollars de revenus par année. Ce profit provient d'une grande diversité de joueurs, allant de jeunes enfants jusqu'à leurs grands-parents avec une population féminine presque aussi importante que celle masculine. Il en résulte donc qu'une grande diversité de jeux et de plates-formes se retrouvent sur le marché.

La grande quantité de jeux qui ont été produits a fait en sorte que les caractéristiques de qualité sont devenues essentielles afin qu'un jeu vidéo puisse se comparer aux jeux déjà existants et ainsi avoir une chance d'être adopté par les consommateurs. Ces attentes du consommateur peuvent se traduire essentiellement par les besoins suivants :

- Exposer un *gameplay* amusant
- Offrir de beaux rendus graphiques et un affichage visuel agréable
- Avoir une composante multi-joueurs
- Optionnellement contenir une narration.

Le *gameplay* d'un jeu représente le facteur de plaisir qu'un usager obtient en jouant à un jeu. Ce facteur est souvent subdivisé en plusieurs catégories comme la diversité et la quantité de niveaux, la difficulté, la réponse des contrôles du jeu, *etc.*

Ces besoins semblent simples *a priori*, mais impliquent beaucoup de travail et imposent des contraintes sévères aux développeurs de jeux vidéo. Afin qu'un jeu puisse offrir un *gameplay* intéressant aux joueurs, une étroite collaboration entre les développeurs et les designers doit être établie. Cela implique aussi d'évaluer plusieurs possibilités en testant sur des prototypes et en effectuant de nombreux cycles de développement du jeu.

En ce qui concerne le rendu graphique du jeu, en plus des designers, c'est aussi avec les artistes que les développeurs doivent s'accorder dans le but de concevoir un moteur de rendu répondant bien aux besoins de leurs créations. Celles-ci doivent être faites en

respectant les contraintes imposées par le matériel où sera déployé le produit. Ainsi, le moteur de rendu doit être facilement extensible afin d’accommoder les nouvelles idées et les nouveaux concepts à intégrer.

Une composante multi-joueurs, tout dépendant si elle implique des parties faites sur le réseau ou non, pourrait nécessiter une grande rigueur de programmation afin d’assurer une stabilité de connexions et empêcher les joueurs de tricher. Ces sujets ne sont pas discutés dans ce mémoire.

La diversité des jeux développés a mené à une classification de ceux-ci. Les principaux genres de jeux sont :

- Action : Jeux rapides demandant souvent de l’habileté de la part des joueurs.
- Stratégie : Jeux à rythme plus lent, exigeant une réflexion plus profonde de la part des joueurs.
- Jeux de rôles : Jeux où la narration de l’histoire occupe une place importante et où les personnages de l’histoire subissent une évolution graduelle en fonction du temps.
- Simulation : Jeux qui tentent de représenter fidèlement des phénomènes réels comme la conduite automobile, des sports, *etc.*
- *Casual* : Jeux simples visant à être utilisés par des joueurs occasionnels. Par exemple, le jeu *Tetris* est considéré comme appartenant à ce genre.

Il existe d’innombrables autres genres et d’hybrides de ces catégories. Notre but ici est simplement de donner une idée de la diversité du contenu des jeux vidéo produits.

2.3 Contraintes de programmation

Le développement de jeux vidéo implique une programmation sous des contraintes sévères afin que le jeu respecte les normes du marché et les attentes des joueurs. Une

contrainte importante pour le développement de jeux vidéo est reliée à l'efficacité algorithmique. En effet, la fluidité d'un jeu est critique face à l'immersion du joueur dans l'univers créé par le jeu de certains genres. Les jeux doivent donc être des programmes efficaces de manière à inviter le joueur à entrer le plus possible dans la narration.

Une autre contrainte importante reliée au développement de jeux est la modularité du code produit. En effet, afin que les efforts placés dans la production d'un jeu vidéo puissent être réutilisés dans les productions subséquentes, des abstractions doivent être faites afin de faciliter la réutilisation de ces dernières. Souvent, pour obtenir des performances adéquates, les développeurs doivent sacrifier la modularité. La conséquence est que chaque jeu représente un énorme travail car peu de code peut être réutilisé.

2.3.1 Fluidité

2.3.1.1 Taux de rafraîchissement

Une autre contrainte dans le développement d'un jeu vidéo est la fluidité de celui-ci. Lorsqu'un film est projeté sur un écran de cinéma, le taux de rafraîchissement de l'image est d'environ 30 trames par seconde. Par contre, les images captées par les caméras contiennent du « flou de mouvement », effet créé par le mouvement s'étant produit durant l'ouverture de l'obturateur pendant un certain temps. Ce flou permet à notre cerveau d'estimer ou de faire l'extrapolation du mouvement dans une trame et donc de croire l'illusion créée par la projection des images. De plus, le faible éclairage présent dans une salle de cinéma aide l'image projetée à rester plus longtemps imprégnée sur la rétine de l'oeil, facilitant ainsi la compréhension du mouvement dans l'image.

Par contre, dans un jeu vidéo, les images rendues à l'écran sont souvent parfaitement nettes et ne contiennent donc pas ce flou de mouvement, ce qui vient réduire la crédibilité de l'illusion faite par la succession des images à l'écran. Il faut donc augmenter le taux de rafraîchissement à environ 60 trames par seconde pour obtenir le niveau de crédibilité d'une projection cinématographique [15].

Toutefois, un taux de rafraîchissement de 60 trames par seconde n'est pas nécessaire

pour tous les jeux. Cela varie grandement avec la nature des jeux. Par exemple, un jeu de stratégie où les joueurs jouent à tour de rôle possède peu d'animations et donc pourrait très probablement être satisfaisant avec un taux de rafraîchissement de 30 trames par seconde ou moins.

Aussi, il est possible que les designers du jeu acceptent de réduire volontairement le taux de rafraîchissement afin d'avoir plus de temps pour rendre une image. Un jeu possédant un taux de rafraîchissement de 60 trames par seconde implique que les images sont rafraîchies toutes les 16 millisecondes, ce qui ne laisse pas beaucoup de temps pour effectuer le calcul du moteur du jeu en plus du rendu de l'image. Heureusement, les cartes vidéo modernes sont capables de faire une bonne partie du rendu laissant ainsi le ou les processeurs principaux libres pour exécuter le moteur du jeu.

Il n'en demeure pas moins que toute la logique du jeu doit être aussi efficace que possible afin que le processus de rendu des trames soit transparent au joueur.

2.3.1.2 Réponse quasi temps réel

Une autre implication de la fluidité requise par un jeu vidéo est le délai de réponse face aux entrées du joueur. Ce délai doit être aussi faible que possible pour donner l'impression que le personnage dans le jeu ne soit qu'une extension de la volonté du joueur. Ainsi le traitement des entrées du joueur se doit aussi d'être très efficace afin de ne pas encombrer le moteur du jeu qui n'a déjà pas beaucoup de temps pour effectuer son travail.

2.3.2 Modularité

En plus de devoir être efficacement implanté, un jeu vidéo moderne se doit d'être aussi modulaire que possible. Pour y arriver, le couplage entre les différentes composantes doit être faible.

Le développement de logiciels favorisant la modularité est une qualité standard en génie logiciel, mais elle vient très bien s'appliquer au développement de jeux vidéo

modernes. En effet, ces derniers sont des projets de très grande envergure impliquant beaucoup de programmeurs. Il en résulte qu'un bon design modulaire permet de bien séparer les tâches à effectuer de manière parallèle par des équipes spécialisées. Si le projet est bien géré, cela pourrait certainement réduire le coût de développement de tels projets.

De plus, une bonne modularité permet de créer des composantes qui sont cohésives et qui permettent donc d'être réutilisées par la suite. Une réutilisation de composantes complexes, tel un moteur de physique, évite de faire un travail supplémentaire non trivial pour chaque nouveau projet. Ainsi, il est clair que l'effort supplémentaire investi pour bien modulariser les composantes d'un jeu seront rentables pour le développement des prochains jeux.

Bien sûr, toute abstraction possède son coût, notamment en coût de performance. Ainsi, il faut bien estimer les endroits où l'utilisation de modules externes n'aura pas de trop grands impacts négatifs de performance.

Il est clair qu'un langage de programmation ayant un bon potentiel d'abstraction, comme c'est généralement le cas pour les langages de haut niveau, facilite cette tâche. Un système orienté objet ou un bon système de module permettrait aussi de bien modulariser les composantes d'un jeu.

2.3.3 Malléabilité

La malléabilité d'un logiciel pourrait se définir comme la facilité avec laquelle ce dernier peut être modifié ou transformé. Cette qualité est importante pour le développement de jeux puisqu'un tel processus implique généralement de faire beaucoup de prototypage. Le code d'un jeu se doit alors d'être très malléable afin d'aider à faire des changements rapidement au moteur du jeu pour tester de nouvelles idées, sans ajouter un trop grand coût pour le faire.

La modularité du jeu aura un effet positif sur la malléabilité en permettant de modifier les mécanismes internes des composantes sans trop affecter le reste du programme.

Par contre, un langage de programmation offrant une bonne puissance d'abstraction serait tout aussi efficace, car il permet d'ajouter de nouvelles abstractions où les besoins de changement apparaissent.

2.4 Conclusion

En conclusion, l'industrie du jeu vidéo et son histoire ont été mises en lumière afin d'exposer l'évolution des jeux vidéo et de motiver l'intérêt de travailler à améliorer le développement de ceux-ci.

Les jeux vidéo sont passés rapidement d'idées farfelues comme ce fut le cas pour le *Cathode Ray Amusement Device* à une industrie complète générant des milliards de dollars annuellement. Il y a un grand intérêt à vouloir diminuer le coût de développement de ceux-ci.

Les coûts de développement de jeux modernes sont de l'ordre des millions de dollars et impliquent des dizaines voire des centaines d'artistes, de programmeurs et de designers. Ainsi, même de toutes petites accélérations d'un cycle de développement pourraient avoir un impact sur le travail de beaucoup de gens et se traduire ainsi en de grandes économies sur les coûts de développement.

Les contraintes impliquées au niveau de la programmation par le développement de jeux vidéo ont été étudiées. Ces dernières se résument par les concepts de fluidité, de modularité et de malléabilité du jeu vidéo. Le respect de la contrainte de fluidité résulte en une bonne immersion pour le joueur, qui est nécessaire afin que le jeu soit considéré comme « jouable ». La modularité facilite le développement en collaboration tout en permettant d'améliorer la réutilisation des composantes communes à plusieurs jeux. L'aspect de la malléabilité du code permet d'accélérer le prototypage du jeu.

Il semble donc que le choix d'un langage permettant de répondre à ces contraintes pourrait certainement faciliter le développement de jeux vidéo. Ainsi, des économies substantielles aux compagnies de l'industrie pourraient être engendrées, rendant plus accessible le développement de jeux aux plus petites compagnies possédant des budgets

moins importants.

CHAPITRE 3

LE LANGAGE SCHEME

Pour la réalisation d'un projet donné, le choix d'un langage de programmation a beaucoup d'influence sur la structure du code écrit. Plusieurs facteurs influencent ces différences. Par exemple, un langage à typage statique résultera en du code qui a une structure adaptée aux besoins du moment, mais qui sera beaucoup moins malléable par la suite [16]. Ainsi, afin d'optimiser les efforts de développement pour un projet, une bonne analyse des besoins est de rigueur afin de choisir un langage répondant le mieux possible à ceux-ci.

En se basant sur les besoins et les contraintes énoncés dans le chapitre 2, nous proposons l'utilisation du langage de programmation Scheme comme outil principal de développement de jeux vidéo. Il s'agit d'un langage de très haut niveau qui semble être un candidat idéal pour permettre un développement efficace de jeux vidéo. En plus de posséder plusieurs implantations très performantes, ce langage offre une grande quantité de concepts de programmation de haut niveau. Parmi ceux-ci, on retrouve entre autres les concepts de la programmation fonctionnelle qui offrent un grand potentiel d'abstraction au niveau de la structure du code, un système de macros évolué qui permet d'augmenter Scheme avec des langages spécifiques au domaine, l'accès aux continuations du calcul qui permettent une manipulation très puissante du flot de contrôle du programme et un système de gestion mémoire automatique.

Ce chapitre donne un bref historique du langage Scheme afin de le mettre en situation et donner un aperçu de son origine. Par la suite, une brève explication des concepts de base du langage est donnée dans le but d'illustrer la simplicité de celui-ci. Plus d'informations sur les bases de Scheme peuvent être obtenues dans la littérature [8, 17]. Finalement, les particularités du langage sont discutées afin de mettre en évidence l'ampleur de celles-ci dans le contexte du développement de jeux vidéo.

3.1 Héritage Lisp

Le langage Scheme est une forme épurée du tout premier langage de programmation de haut niveau, le langage Lisp. Ainsi, pour illustrer les origines de Scheme, une courte histoire du langage Lisp est présentée. Par la suite, les langages de programmation qui furent inspirés du langage Scheme sont mentionnés en expliquant brièvement leur racine commune et leurs principales disparités.

3.1.1 Histoire de Lisp

Le langage Lisp est le tout premier langage de programmation de haut niveau qui fut conçu et est le deuxième plus vieux langage de programmation toujours utilisé. Sa première implantation fut réalisée en 1958 dans un laboratoire d'intelligence artificielle du MIT (*Massachusetts Institute of Technology*) sous la direction de John McCarthy. Le langage original introduisait déjà des concepts fondamentaux comme l'utilisation de listes pour représenter du code, la programmation fonctionnelle, ainsi que le concept de gestion de mémoire automatique. Un article décrivant le langage Lisp comme un langage de programmation et de formalisme de définition de fonctions récursives fut publié ultérieurement [18].

Le langage fut aussitôt adopté par la communauté de recherche en intelligence artificielle, puisque l'utilisation d'un traitement symbolique permettait de développer de nouveaux algorithmes dans ce domaine.

Lisp évolua par la suite en de nombreux dialectes différents, chacun apportant ses propres extensions au langage. Vers le début des années 1980, un comité d'unification de ces multiples dialectes fut formé pour donner naissance au standard Common Lisp. En 1984, Guy Steele publia un livre [19] donnant une description détaillée du standard établi par la communauté Lisp. Il en résulte donc une spécification très complexe du langage qui s'est trouvée gonflée par les multiples dialectes lui ayant donné naissance. Ainsi, Common Lisp possède beaucoup de bagage, et il devient difficile pour un programmeur de maîtriser tous les concepts de ce langage.

3.1.2 Avènement de Scheme

Peu avant la standardisation de Common Lisp, Guy Steele et Gerald J. Sussman concurent la spécification d'un nouveau dialecte du langage Lisp appelé Scheme [20]. L'origine du développement de Scheme provient de la tentative de compréhension d'un formalisme modulaire d'agents en intelligence artificielle [21, 22]. Ils ont donc bâti un interprète minimaliste basé sur Lisp afin de comprendre ce formalisme qui s'exprimait mal avec les langages disponibles à l'époque. Ce petit interprète supportait ainsi une portée statique appropriée pour la description d'acteurs, divergeant ainsi de Lisp en remplaçant de simples fonctions récursives par les fermetures du lambda calcul. Les deux auteurs furent très surpris de constater que le langage expérimental créé répondait parfaitement à leurs besoins et s'avérait beaucoup plus simple que ce qu'ils avaient anticipé.

Scheme devint ainsi populaire et est présentement l'un des deux dialectes les plus utilisés de Lisp, en concurrence avec Common Lisp. Même après avoir subi cinq révisions, le langage Scheme demeure très simple avec un standard ne comptant qu'environ une trentaine de pages [8]. Il en résulte qu'on ne retrouve actuellement pas moins d'une cinquantaine d'implantations disponibles, dont une douzaine sont considérées comme étant des implantations majeures. Celles-ci offrent non seulement plusieurs extensions du langage, mais aussi de bonnes performances grâce à l'implantation d'optimisations de compilation avancées.

Le langage Scheme impose une philosophie de minimalisme à ses programmeurs en ne leur fournissant que la base dans le standard. Cette philosophie est souvent appréciée car elle mène à utiliser ce qui est vraiment nécessaire et à faire soi-même ce qui n'est pas disponible, créant ainsi une bonne autonomie chez les programmeurs Scheme. De plus, le langage est accompagné d'une centaine de bibliothèques externes de traitement générique dénommées *Scheme Request For Implementation* (SRFI [23]).

3.1.3 Langages inspirés de Scheme

Plusieurs langages de programmation récents furent inspirés des langages Lisp. Pour la plupart, il s'agit de langages de script comme JavaScript [24], Ruby [25] ou Perl [26]. Ces langages s'inspirent notamment de l'aspect dynamique de Lisp afin d'être interprétés. Ils utilisent aussi une gestion de mémoire automatique afin de simplifier leur utilisation. Finalement, certains d'entre eux offrent des fonctions en tant que données de première classe, permettant ainsi l'utilisation de la programmation fonctionnelle.

Par contre, ces langages demeurent toujours des sous-ensembles de Scheme en terme de fonctionnalités. Ils sont devenus populaires pour leur simplicité d'utilisation, mais un bon apprentissage de Scheme pourrait certainement être plus bénéfique à long terme. Entre autres, Scheme est un langage interprété et compilé, permettant de développer des applications très efficaces. Il en résulte que le code peut être développé de manière dynamique en utilisant une *Read Eval Print Loop* (*REPL*) afin de tester et corriger un programme développé. La présence d'une *REPL* permet aussi un mode de récupération d'erreurs très dynamique et donnant accès à l'évaluation d'expressions Scheme lorsqu'une erreur se produit dans une telle boucle.

3.2 Introduction au langage

Tout en bénéficiant du riche héritage des langages Lisp, Scheme demeure un langage épuré et très simple. Cette section vise à présenter les bases du langage Scheme et de son implantation par le système Gambit-C [9], afin de familiariser le lecteur avec la syntaxe du langage ainsi que les fonctionnalités de base et les extensions au langage fournies par Gambit-C.

3.2.1 Syntaxe

Les langages de la famille Lisp adoptent une notation préfixe parenthésée. Il en résulte que la syntaxe du langage est consistante et simple. La figure 3.1 illustre un exemple simple de calcul arithmétique.

```
(+ (expt 2 5) (sin (/ 3.141592653589793 2))) => 33.
```

FIGURE 3.1 – Exemple simple d’appels de fonctions en Scheme

Cette syntaxe apporte plusieurs bénéfices au langage. Puisque chaque appel de fonction (ou de forme spéciale) est parenthésé, il en résulte qu’un programme Scheme est très structuré et peut être manipulé efficacement en utilisant des outils de développement performants tels *VI* ou *Gnu Emacs*. De plus, Scheme utilise la même syntaxe pour les données d’un programme que pour le programme. Il devient ainsi possible de concevoir des programmes qui créent des programmes Scheme. C’est ce qui constitue la base du système de macros de Scheme. On dit ainsi qu’un programme Scheme est représenté par des expressions symboliques, fréquemment appelées S-expression. La distinction entre une donnée et les instructions d’un programme se fait par l’entremise des formes spéciales `quote` et `quasiquote` qui spécifient respectivement que l’expression doit être entièrement ou en partie traitée comme une donnée et donc, ne doit pas être évaluée comme un programme. La figure 3.2 donne deux exemples d’utilisation de ces formes spéciales. Le premier exemple illustre l’utilisation de `quote` qui fait en sorte que la liste spécifiée est considérée comme une donnée et non un appel de la fonction `abc`. Le deuxième exemple utilise la forme `quasiquote` qui permet de spécifier des données dont certaines parties peuvent être évaluées à l’aide de la forme spéciale `unquote`. Les formes spéciales `quote`, `quasiquote` et `unquote` peuvent être aussi utilisées *via* les caractères respectifs `’`, `‘` et `,.`

```
(+ 1 (list-ref '(abc 123 "toto") 1)) => 124
(list-ref `(abc ,(+ 122 1) "toto") 1) => (abc 123 "toto")
```

FIGURE 3.2 – Exemples d’utilisation de listes sous forme de S-expression

3.2.2 Aperçu des fonctionnalités requises par le standard

Le standard du langage Scheme [8] est très minimaliste, mais possède toutes les fonctionnalités de base requises par un langage de haut niveau comme Scheme. On y

spécifie que Scheme est un langage typé dynamiquement, *i.e.* que les vérifications de types sont faites durant l'exécution d'un programme, et non durant la compilation. Cette particularité fait en sorte que Scheme est un langage qui peut être interprété ou compilé. Les types de base du langage sont tous disjoints entre eux, *i.e.* aucun objet ne peut appartenir à plus d'un de ces types. Ces types sont vérifiés par les prédicats `boolean?`, `pair?`, `symbol?`, `number?`, `char?`, `string?`, `vector?`, `port?`, `procedure?` et `null?`. Chacun de ces types est documenté et possède des fonctions standards de création et d'utilisation. Entre autres, la forme spéciale `lambda` permet la création de fermetures telles qu'elles existent dans le lambda calcul. La figure 3.3 illustre l'utilisation des quelques-uns parmi ces types de base.

```
((lambda (x y) (+ (car x) (vector-ref y 2)))
 '(1 2 3)
 '#(4 5 6)) => 7

(substring (list->string '("#\a #\l #\l #\o #\!))
 0
 3) => "all"
```

FIGURE 3.3 – Exemples d'utilisation des types de base de Scheme

3.3 Programmation fonctionnelle

Le langage Scheme est un langage fonctionnel. Cette famille de langages de programmation est caractérisée par une programmation centrée sur l'appel de fonctions dites avec une transparence référentielle, où la valeur de retour d'une fonction dépend uniquement des paramètres et non de l'état du système. Ce type de programmation fait contraste à la programmation impérative, qui est centrée sur la mutation de l'état d'un programme. Les programmes écrits en utilisant ces langages sont généralement beaucoup plus faciles à comprendre, ce qui encourage un grand nombre de chercheurs à contribuer au développement et l'utilisation de tels langages dans le milieu commercial.

3.3.1 Programmation fonctionnelle pure

Une programmation dite purement fonctionnelle respecte de manière stricte le paradigme de la programmation fonctionnelle, *i.e.* que les fonctions sont des entités mathématiques qui ne sont pas affectées par l'état du système. L'origine de ce type de programmation se situe avant même l'apparition des premiers ordinateurs. C'est dans les années 1936 qu'Alonzo Church a introduit le formalisme mathématique de définition de fonctions mathématiques dénommé lambda calcul non typé [27]. Ce dernier introduit un formalisme de définition de fonctions qui sert de base aux langages de programmation fonctionnels. La figure 3.4 illustre la définition de la fonction `flip`. Cette dernière reçoit en paramètre une fonction f et deux valeurs (y et x) et s'occupe d'appliquer la fonction reçue aux paramètres dans l'ordre inverse.

$$\lambda f \rightarrow \lambda y \rightarrow \lambda x \rightarrow fxy$$

FIGURE 3.4 – Définition de la fonction `flip` en lambda calcul non typé

Ces fonctions sont considérées pures car elles sont exemptes de mutations de l'état du calcul. Ainsi, le résultat de l'application d'une fonction (appelée β -reduction en lambda calcul) ne dépend que des paramètres de ces dernières.

Une caractéristique importante présente dans le lambda calcul est la présence de fonctions de première classe, *i.e.* que les fonctions sont des objets appartenant aux valeurs du langage. Elles peuvent donc être créées et manipulées comme toutes autres valeurs. Ces fonctions sont aussi appelées fermetures, car elles implantent une portée statique de variables, *i.e.* qu'elles conservent l'environnement d'exécution dans lequel elles se trouvaient lors de leur création. Cette capacité de retenir l'environnement est très puissante et permet l'utilisation de fermeture comme structure de données. Par exemple, la figure 3.5 illustre l'implantation de paires similaires à celles retrouvées dans Scheme en lambda calcul.

Comme déjà mentionné dans la section 3.1, le lambda calcul a donné naissance à une

$$\begin{aligned}
cons &= \lambda car \rightarrow \lambda cdr \rightarrow \lambda f \rightarrow fxy \\
car &= \lambda pair \rightarrow pair(\lambda car \rightarrow \lambda cdr \rightarrow car) \\
cdr &= \lambda pair \rightarrow pair(\lambda car \rightarrow \lambda cdr \rightarrow cdr)
\end{aligned}$$

$$car (cons\ 1\ 2) \rightsquigarrow 1$$

FIGURE 3.5 – Implantation de paires en utilisant des fermetures du lambda calcul

famille de langages dont la racine est le langage Lisp. Dans cette famille, les langages ML [28] et Haskell [29] sont parmi les plus connus de ceux qui adhèrent fidèlement au modèle de calcul d’Alonzo Church. Le langage Scheme est lui aussi un langage fonctionnel, mais il n’est pas considéré pur puisque les mutations de l’état du programme sont permises. En effet, des opérateurs tels que `set!`, `set-car!` ou encore `vector-set!` permettent la modification de liaison ou du contenu de cellules mémoires. Ainsi, Scheme partage le meilleur des deux mondes en laissant la liberté au programmeur de choisir le paradigme de programmation qui lui apparaît le plus approprié. Par contre, un abus de mutations est considéré comme étant de très mauvais style en Scheme.

3.3.2 Programmation fonctionnelle en Scheme

Le lambda calcul ouvre la porte à la création d’abstractions très intéressantes : les fonctions d’ordre supérieur. Celles-ci reçoivent des fonctions en paramètre, les manipulent et en retournent de nouvelles. La figure 3.6 illustre un exemple d’écriture et d’utilisation de fonctions d’ordre supérieur. La première fonction définie prend une fonction en paramètre et applique cette fonction sur chacun des éléments de la liste passée comme deuxième paramètre. La deuxième fonction est une version Scheme de `flip` présentée dans la figure 3.4. L’exemple utilise ensuite ces deux dernières afin de soustraire 1 à chaque élément de la liste spécifiée.

L’utilisation de fonctions d’ordre supérieur permet l’abstraction d’algorithmes de manière très générique. Une autre fonction d’ordre supérieur très répandue est la fonc-


```

(define (map f lst)
  (if (pair? lst)
      (cons (f (car lst)) (map f (cdr lst)))
      '()))

(define (flip f y)
  (lambda (x) (f x y)))

(map (flip - 1) '(1 2 3 4 5)) => (0 1 2 3 4)

```

FIGURE 3.6 – Exemple de création et d'utilisation de fonctions d'ordre supérieur.

tion `foldl` qui permet d'utiliser chacun des éléments d'une liste pour calculer un résultat unique. La figure 3.7 illustre l'utilisation de cette fonction pour faire la somme et le produit des nombres dans une liste.

```

(define (foldl f acc lst)
  (if (not (pair? lst))
      acc
      (foldl f (f acc (car lst)) (cdr lst))))

(foldl + 0 '(1 2 3 4 5)) => 15

(foldl * 1 '(1 2 3 4 5)) => 120

```

FIGURE 3.7 – Fonction d'ordre supérieur `foldl`.

Ainsi, l'utilisation de fonctions en tant que données de première classe permet la création de puissantes abstractions algorithmiques. De plus, l'utilisation des paradigmes de programmation fonctionnelle facilite beaucoup la modularisation du logiciel. En effet, puisque peu ou pas de mutations d'état sont utilisées, il devient alors facile de combiner différents modules, sans avoir d'interférences dans l'état du programme, ce qui n'est pas nécessairement le cas avec des modules impératifs dépendant directement de l'état du programme pour leur exécution.

De plus, il est possible en Scheme de créer des fermetures n'attendant aucun argument. Ces dernières sont généralement appelées *thunk* de manière historique. Elle sont utilisées afin de retarder l'évaluation de code dans le but de faire de l'évaluation paresseuse ou d'encapsuler un comportement. La figure 3.8 illustre l'utilisation d'un

thunk avec le système de *threads* de Gambit-C afin de suspendre l'évaluation du corps du processus léger créé jusqu'à ce que ce dernier soit démarré.

```
(let ((thunk (lambda () (display 'B))))
  (display 'A)
  (thread-start! (make-thread thunk))) => AB
```

FIGURE 3.8 – Utilisation d'un *thunk* dans le but de retarder l'évaluation de code

3.4 Macros

Les macros d'un langage de programmation sont des systèmes utilisés pour faire des abstractions qui sont résolues durant la compilation d'un fichier source. Dans sa forme la plus élémentaire, un système de macro peut être implanté sous la forme d'un pré-processeur de code qui ne fait que du remplacement de patrons par une expansion textuelle directe. C'est ce genre de système de macros que l'on retrouve entre autres dans le langage C. Ce type de macro, quoique très simple, permet de faire déjà beaucoup d'abstractions. La figure 3.9 illustre un exemple de macros C. Dans cet exemple, on définit la macro `swap` qui interchange la valeur de deux variables entre elles. Le résultat de l'expansion sera donné par la substitution de l'appel de la macro par son corps où les occurrences des paramètres sont intégralement substitués par les paramètres actuels, soient `a` et `b`.

```
#define swap(x,y) do { int tmp; tmp = x; x = y; y = tmp; } while(0)

int main(){
  int a = 1; int b = 2;
  swap(a,b);
  return a - (b<<1);
}
```

FIGURE 3.9 – Exemple de macro textuelle en C

De telles macros comportent beaucoup de problèmes potentiels. Non seulement,

elles sont limitées à faire des remplacements syntaxiques très primitifs, mais aussi elles sont souvent la source de problèmes reliés au changement du contexte syntaxique des arguments de la macro ou à une expansion dans un contexte qui n'était pas prévu. Ces problèmes sont difficiles à trouver car ils surviennent uniquement lorsque le programme est compilé, bien après l'expansion des macros. Il en résulte que le code compilé ne contient aucune trace des appels macros et donc le compilateur donne souvent des erreurs qui ne sont pas reliées avec la vraie source du problème, soit la définition de la macro en question.

L'utilisation d'expressions symboliques pour représenter de la même manière les données d'un programme et le code source de celui-ci en Scheme permet l'implantation d'un système de macros très puissant, dit procédural. Plusieurs de ces systèmes d'expansion de macros sont disponibles en Scheme. Le plus simple d'entre eux est celui qui provient directement de Lisp, soit la forme spéciale **define-macro**. Avec cet expanseur, une macro Scheme est une fonction dont les paramètres *ne sont pas évalués* avant d'être passés au corps de la fonction. Il en résulte ainsi que le code étant en position d'argument est considéré par la macro comme étant des données Scheme (listes, symboles, nombres, *etc.*). Puisque la macro Scheme est une fonction, elle possède toute la puissance du langage à sa disposition. La S-expression retournée par la macro sera le résultat de l'expansion de celle-ci et sera donc évaluée durant l'exécution du programme à l'endroit où l'appel de la macro se retrouvait. Un exemple simple de macro tirant profit du fait que l'argument est passé par nom est illustré dans la figure 3.10. Cette macro très simple utilise le passage par nom pour permettre d'incrémenter la valeur d'une variable, ce qui ne serait pas possible autrement en Scheme puisque le passage d'argument se fait normalement par valeur.

Un autre exemple de macro utilisant le passage par nom est donné dans la figure 3.11. Cette macro permet d'augmenter le langage Scheme d'une nouvelle forme spéciale qui correspond à une forme simple d'itération comme l'on retrouve souvent dans les

```
(define-macro (inc! var) `(set! ,var (+ ,var 1)))

(let ((x 1))
  (inc! x)
  x) => 2
```

FIGURE 3.10 – Exemple simple de macro utilisant le passage par nom à profit

langages impératifs. L’expansion de l’appel à cette macro illustrée dans la figure 3.11 est présentée dans la figure 3.12.

```
(define-macro (for var init limit . body)
  `(let loop ((,var ,init))
    (if (< ,var ,limit)
        (begin ,@body
                 (loop (+ ,var 1)))))

(let ((v (make-vector 5)))
  (for x 0 5 (vector-set! v x x))
  v) => #(0 1 2 3 4)
```

FIGURE 3.11 – Exemple de macro ajoutant une nouvelle forme spéciale au langage

```
(let ((v (make-vector 5)))
  (let loop ((x 0))
    (if (< x 5) (begin (vector-set! v x x)
                       (loop (+ x 1)))))
  v)
```

FIGURE 3.12 – Résultat de l’expansion macro de l’appel présenté dans la figure 3.11

Ces macros demeurent très simples et se rapprochent de ce qui peut être exprimé à l’aide de macros C. Toutefois, la macro `for` comporte un problème. Celui-ci est relié à l’introduction de la variable `loop`. L’ajout de cette variable peut causer le problème connu sous le nom de *Capture de nom*. En effet, la macro n’aura pas le comportement escompté si l’un de ses paramètres possède une référence à une variable déjà existante nommée `loop`. Le problème de capture de variable était tout aussi présent dans les macros de C et constituait pour ce type de macro un problème insolvable. Par contre, il est possible d’éviter ce problème en Scheme en générant un nom de variable pour `loop` qui est assurément unique. L’implantation Gambit-C offre la fonction `gensym` pour ces

besoins.

Le problème de capture de nom est associé au phénomène appelé hygiène des macros. Ce problème peut être résolu manuellement comme suggéré, mais il existe aussi d'autres expanseurs de macros assurant l'hygiène des macros, dont entre autres, la forme spéciale `syntax-rules` [30]. La réécriture de la macro `for` en utilisant un système de macros hygiénique est donnée dans la figure 3.13. Cette figure donne aussi un exemple d'utilisation de la macro mettant à l'épreuve l'hygiène de celle-ci. On constate que malgré le fait qu'aucun effort n'a dû être mis en place pour éviter les collisions de nom pour la variable `loop` introduite, la macro réalise le comportement attendu dans des conditions potentiellement problématiques.

```
(define-syntax for
  (syntax-rules ()
    ((for var init limit body ...)
     (let loop ((var init))
       (if (< var limit)
           (begin body ...
                   (loop (+ var 1)))))))

(let ((v (make-vector 5)))
  (for loop 0 5 (vector-set! v loop loop))
  v) => #(0 1 2 3 4)
```

FIGURE 3.13 – Macro `for` hygiénique

Bien que l'expansion de macro faite avec `define-macro` puisse impliquer des collisions de variable, elle permet une plus grande liberté d'écriture. En effet, parfois la collision de nom peut être intentionnelle. Aussi, le problème est résolu avec l'utilisation de symboles uniques.

Un exemple plus convaincant d'une macro Scheme utilisant la puissance de calcul du langage est illustré dans la figure 3.14. Cette macro résulte en la création d'un masque binaire ayant les bits des puissances de deux spécifiées en argument. Ainsi, l'expansion de l'appel `(mask 1 3 8)` est directement 266. Il est intéressant de noter qu'un appel à cette macro tel que `(mask 1 (+ 1 2))`, est une erreur. En effet, la macro s'attend à

avoir un nombre en paramètre, mais elle reçoit plutôt dans ce cas-ci la liste `(+ 1 2)`, puisque les arguments de macros ne sont pas évalués.

```
(define-macro (mask . bits)
  (define (curry2 f x) (lambda (y) (f x y)))
  (apply + (map (curry2 expt 2) bits)))

(number->string (mask 1 3 8) 2) => "100001010"
```

FIGURE 3.14 – Exemple plus avancé de macro Scheme

Le calcul de la somme des puissances de deux se fait donc durant la compilation du programme, donnant accès à des optimisations intéressantes, sans toutefois perdre au niveau de l'élégance du code. Ce genre de macros mènent directement vers la construction de langages spécifiques au domaine (*Domain Specific Languages* ou simplement *DSL*). Ces derniers consistent en l'ajout de nouvelles constructions à Scheme permettant d'exprimer des idées reliées à un concept précis de manière élégante, sans perte de performance. La macro `mask` pourrait constituer le pilier de base d'un *DSL* effectuant la création et la manipulation de masques binaires. La figure 3.15 illustre une deuxième macro effectuant durant l'expansion macro la combinaison de masques binaires. On obtient un petit langage pour la déclaration efficace de tels masques.

```
(define-macro (mask-or . masks)
  (apply bitwise-ior (map eval masks)))

(mask-or (mask 1 3) (mask 8)) => 266
```

FIGURE 3.15 – Macro venant augmenter le petit *DSL* portant sur la manipulation de masques binaires

On peut maintenant imaginer de tels langages spécifiques au domaine pour n'importe quel problème associé aux jeux vidéo. Par exemple, une macro peut être écrite pour la description détaillée de modèles tridimensionnels et une autre pour la combinaison de ces derniers. Ainsi, ces macros pourraient permettre de décrire le modèle d'une manière simple pour les artistes, en utilisant une syntaxe près d'un langage naturel.

Durant l'expansion, celles-ci transformeraient ces déclarations et combinaisons en des structures de données efficaces tant en utilisation mémoire qu'en rapidité d'accès et de manipulation. Un autre exemple pourrait être la création d'un système de programmation orientée objets à l'aide de macros Scheme. La puissance de calcul disponible durant l'expansion pourrait ainsi servir à la création des fonctions auxiliaires d'accès aux membres, à la détermination des membres hérités, *etc.*

En conclusion, les macros Scheme donnent accès à une très grande puissance expressive en permettant l'utilisation de la puissance complète du langage durant l'expansion de ces dernières. Cette puissance de calcul et cette simplicité proviennent du fait que le code source et les données utilisent la même syntaxe, soit les S-expressions. Ainsi, en utilisant ce système de macros, il est possible d'ajouter de nouvelles formes spéciales au langage et créer des langages spécifiques au domaine (*Domain Specific Languages*).

3.5 Continuations

La forme d'écriture de code appelée CPS (*Continuation Passing Style*) est un style d'écriture utilisé dans les langages fonctionnels qui permet d'écrire la suite du calcul d'une fonction explicitement. La figure 3.16 illustre une implantation simple de la fonction factorielle. La suite du calcul effectué après un appel à la fonction **fact** est implicite à l'endroit où celui-ci se retrouve. Dans l'exemple d'appel à **fact** présenté dans cette figure, on constate que la suite de ce calcul sera l'addition de 1 au résultat obtenu.

```
(define (fact n)
  (if (< n 2)
      1
      (* n (fact (- n 1)))))

(+ (fact 10) 1) => 3628801
```

FIGURE 3.16 – Implantation simple de la fonction factorielle

Le corps de la fonction **fact** présentée possède deux types d'appels de fonctions, soient les appels terminaux et non-terminaux. La distinction entre ces deux types d'ap-

pels est la suivante : un appel terminal est un appel qui termine l'exécution d'une fonction et un appel non-terminal est tel que le résultat de cet appel nécessitera du traitement ultérieur dans la fonction avant que celle-ci termine. Ainsi, seul l'appel à la fonction `*` est terminal dans la fonction `fact`. Le standard du langage Scheme requiert l'implantation de l'optimisation d'appels terminaux. Cette dernière fait en sorte que le résultat d'un appel terminal est passé directement à la continuation de l'appel à la fonction en question. Le résultat d'un appel terminal ne sera donc pas attendu par la fonction qui le contient. Il est ainsi possible d'effectuer des itérations à l'aide d'appels terminaux de fonctions.

La transformation de cette fonction en style CPS est donnée dans la figure 3.17. Dans cette nouvelle version, la suite de l'exécution d'un appel est explicite. Il s'agit d'un appel à la continuation passée en paramètre, nommée `k`.

```
(define (minusk n1 n2 k) (k (- n1 n2)))

(define (timesk n1 n2 k) (k (* n1 n2)))

(define (factk n k)
  (if (< n 2)
      (k 1)
      (minusk n 1 (lambda (r)
                    (factk r (lambda (r2)
                              (timesk n r2 k)))))))

(factk 10 (lambda (x) x)) => 3628800
```

FIGURE 3.17 – Forme CPS de la fonction factorielle

Une propriété intéressante de cette transformation de code est reliée au fait que tous les appels de fonctions deviennent des appels terminaux. La transformation CPS est souvent utilisée par les implantations de Scheme car elle facilite l'implantation de l'optimisation d'appels terminaux. Ce style de programmation permet aussi d'utiliser de manière élégante la puissance d'abstraction des langages fonctionnels en divisant le calcul en petites parties interchangeables et modulaires.

Le langage Scheme offre la possibilité de réifier automatiquement la continuation

d'un calcul donné par l'entremise de la forme spéciale `call-with-current-continuation` souvent nommée aussi `call/cc`. Cette dernière s'attend à recevoir une fermeture en argument, celle-ci n'attendant qu'un seul paramètre : la continuation réifiée du calcul au point d'appel de `call/cc`. La figure 3.18 donne un exemple de la réification d'un calcul simple. La mutation de la variable `k` est utilisée afin de permettre une sauvegarde de la continuation dans le but de la réutiliser plus tard.

```
(define k #f)
(define f (lambda (x y) (* x y)))

(+ 3 (- (call/cc (lambda (cont) (set! k cont) (f 5 6))) 2)) => 31
(k 10) => 11
```

FIGURE 3.18 – Réification d'une continuation à l'aide `call/cc`

Ainsi, la sauvegarde d'une continuation permet de conserver l'état du calcul et la réutilisation de celui-ci.

3.5.1 Exemples d'utilisations

L'utilisation de continuations ajoute beaucoup de puissance au niveau de la programmation en Scheme, si elle est habilement exécutée. Cette section présente quelques utilisations intéressantes permettant d'entrevoir les possibilités qu'offre cet aspect de Scheme.

3.5.1.1 Échappement au flot de contrôle

Un exemple simple d'utilisation de continuation est l'échappement au contrôle d'une fonction. Un programmeur a toujours le contrôle du programme qu'il écrit, mais il peut y avoir des situations qui exigent l'utilisation de fonctions externes dont on ne peut pas modifier le contenu. De telles situations peuvent mener à des inefficacités algorithmiques. La figure 3.19 illustre une utilisation de la fonction d'ordre supérieur `foldl` présentée dans la figure 3.7 afin d'implanter un algorithme permettant de déterminer s'il existe au moins une valeur supérieure à 10 dans une liste. En utilisant la fonction-

```

(trace foldl)

(define (exists-greater-than val lst)
  (foldl (lambda (false x) (if (> x val) x false))
    #f
    lst))

(exists-greater-than 100 (list 1 102 3 2 7 12))

|>(foldl proc #f '(1 102 3 2 7 12))
|>(foldl proc #f '(102 3 2 7 12))
|>(foldl proc 102 '(3 2 7 12))
|>(foldl proc 102 '(2 7 12))
|>(foldl proc 102 '(7 12))
|>(foldl proc 102 '(12))
|>(foldl proc 102 '())
|102
102

```

FIGURE 3.19 – Utilisation non efficace de `foldl`

nalité de traces qu’offre Gambit-C, on constate que malgré le fait que l’utilisation d’une fonction d’ordre supérieur permet d’implanter cet algorithme, il en résulte du code inefficace, puisque l’appel à `foldl` passera sur tous les éléments de la liste, même après avoir trouvé une valeur supérieure à 10.

On doit donc ré-écrire le méta-comportement offert par `foldl` afin d’arrêter la recherche aussitôt qu’une valeur supérieure à 10 est trouvée. Cela vient à l’encontre du principe d’abstraction et mène à une maintenance difficile du code.

Les continuations nous offrent par contre la possibilité d’utiliser la méta-fonction `foldl`, sans perte de performance. La figure 3.20 indique comment s’y prendre. En réifiant la continuation au début du calcul de la fonction `exists-greater-than`, on obtient une fonction qui permet le branchement direct vers la suite du calcul à l’appel de `exists-greater-than`. On peut ainsi utiliser cette dernière afin d’échapper le corps de la fonction lorsqu’une valeur clé est trouvée.

```

(trace foldl)

(define (exists-greater-than val lst)
  (call/cc
    (lambda (return)
      (foldl (lambda (retval x) (if (> x val)
                                   (return x)
                                   retval))
              #f
              lst))))

(exists-greater-than 100 (list 1 102 3 2 7 12))

|>(foldl proc #f '(1 102 3 2 7 12))
|>(foldl proc #f '(102 3 2 7 12))
102

```

FIGURE 3.20 – Utilisation d’une continuation pour échapper au flot de contrôle

3.5.1.2 Recherche par retour arrière

Un autre exemple d’utilisation intéressant de continuations est la création d’un système de recherche par retour arrière (*backtracking*). La figure 3.21 illustre un exemple simple de recherche d’un triplet de nombres entiers tel que $x^2 = y^2 + z^2$.

```

(define fail (lambda () (error "can't backtrack")))

(define (in-range a b) (call/cc (lambda (cont) (enumerate a b cont))))

(define (enumerate a b cont)
  (if (> a b)
      (fail)
      (let ((save fail))
        (set! fail (lambda () (set! fail save)
                              (enumerate (+ a 1) b cont)))
        (cont a))))

(let ((x (in-range 1 9))
      (y (in-range 1 9))
      (z (in-range 1 9)))
  (if (= (* x x) (+ (* y y) (* z z)))
      (list x y z)
      (fail))) => (5 3 4)

```

FIGURE 3.21 – Exemple de recherche par retour arrière

Dans cet exemple, la fonction `fail` représente une liste chaînée de retour arrières permettant de revenir à une étape de décision précédente et de poursuivre le calcul en prenant une nouvelle décision à cet endroit. Cet exemple illustre bien l'élégance qui peut résulter de l'utilisation judicieuse de la forme spéciale `call/cc`.

3.5.2 Conclusion

Ainsi, la réification de continuations implicites ou encore l'utilisation d'un style CPS utilisant des continuations explicites apporte beaucoup de puissance au langage Scheme. L'utilisation de continuations permet la sauvegarde de l'état d'un calcul en cours dans une fermeture et permet sa réutilisation ultérieure. On peut ainsi créer sans trop d'efforts des systèmes d'exceptions, de retour arrières ou même de coroutines, systèmes qui seraient tous pertinents à utiliser dans le cadre de la programmation de jeux vidéo.

3.6 Gestion mémoire automatique

L'écriture de code effectuant la gestion de la mémoire d'un programme a toujours été une partie délicate du développement d'un logiciel. Non seulement, il est très facile de faire des erreurs, mais ces erreurs sont très difficiles à détecter et à trouver. Les problèmes liés à une mauvaise gestion de la mémoire sont classifiés en deux catégories : des fuites de mémoire ou des pointeurs fous. Les fuites de mémoire sont causées par la rétention de mémoire par des objets qui ne sont plus utilisés par le programme. Les pointeurs fous sont des pointeurs toujours accessibles par le programme, mais dont l'espace mémoire correspondant a été récupéré par le système d'exploitation. Ces erreurs sont souvent les conséquences d'un problème plus difficile, soit le problème de modularité. En effet, la durée de vie d'un objet en mémoire est normalement facile à déterminer à petite échelle. Par contre, lorsque le code est écrit de manière modulaire, il est difficile de prévoir toutes les utilisations des objets et souvent la responsabilité de la désallocation des objets est difficile à déterminer.

Ces erreurs sont très problématiques. C'est sans doute pour cette raison que déjà très tôt, des systèmes de gestion automatique de mémoire ont été développés. Les premiers systèmes de gestion de mémoire automatique, appelés aussi ramasse miettes, apparurent avec le langage Lisp. Le principe de base est simple, il s'agit de parcourir la mémoire utilisée par le programme en partant des *racines* de celui-ci. Les racines comprennent tous les liens aux données accessibles, comme les variables globales ou locales d'une fonction en cours d'exécution. Lorsque la mémoire est parcourue, on garde trace des zones accessibles. Après avoir parcouru toute la mémoire utilisée, il ne suffit que de récupérer automatiquement les zones qui ne sont plus accessibles par le programme.

Bien sûr, la gestion de mémoire automatique se fait avec un certain coût en temps de calcul. La répartition de ce temps dépend grandement de la technique de gestion utilisée. Aussi, les algorithmes de gestion de mémoire se doivent d'être *conservateurs* dans la désallocation d'objets puisque la détermination de la vivacité d'un objet est, dans la généralité du problème, indécidable. Ainsi, il peut y avoir des cas où le système de gestion pense qu'un pointeur est toujours accessible, alors qu'en réalité, il ne sera plus jamais utilisé par le programme, créant ainsi des fuites de mémoire. Ces fuites sont toutefois beaucoup plus difficiles à créer que celles obtenues lorsqu'une gestion manuelle de mémoire est utilisée.

3.6.1 Survol des techniques

Plusieurs techniques de gestion de mémoire automatique ont été développées au cours des années. Cette section effectue un bref survol de ces différentes techniques. Plusieurs articles effectuant une revue des techniques existantes sont disponibles [31,32].

3.6.1.1 *Mark and Sweep*

Une des techniques les plus simples de gestion de mémoire automatique est le *Mark and Sweep* [18]. Cette technique consiste à traverser la mémoire à partir des racines et à marquer chacun des espaces mémoires rencontrés lors du parcours. Ensuite, tous

les espaces mémoires alloués sont parcourus et tous ceux qui ne sont pas marqués sont récupérés. Cette technique est très simple, mais peut mener à des problèmes de fragmentations de la mémoire. De plus, elle requiert de traverser toute la mémoire utilisée par deux fois, impliquant donc une pause considérable dans l'exécution du programme.

3.6.1.2 *Stop and Copy*

Une autre technique classique se nomme *Stop and Copy* [33]. De manière similaire au *Mark and Sweep*, cette technique parcourt les zones mémoires utilisées à partir des racines. Par contre, le *Stop and Copy* doit séparer la mémoire disponible en deux zones distinctes, l'une dans laquelle les allocations sont effectuées et l'autre demeure réservée. Lorsque le ramasse miettes atteint un objet en mémoire, il le copie vers la nouvelle zone réservée et laisse une indication de sa nouvelle position dans son ancien emplacement. Toutes les zones mémoires sont ainsi copiées et compactées. Donc, les allocations subséquentes sont par la suite effectuées dans la nouvelle zone mémoire, résultant en une désallocation implicite des espaces inutilisés dans la zone précédente. Cette technique est légèrement plus efficace que le *Mark and Sweep* traditionnel, mais nécessite le double d'espace pour faire la gestion de la mémoire. Le *Stop and Copy* est utilisé dans l'implantation Scheme Gambit-C.

3.6.1.3 Gestionnaires générationnels

Une amélioration notable de la technique du *Mark and Sweep* a donné naissance aux ramasse miettes générationnels [34]. Brièvement, cette technique divise la mémoire en plusieurs générations. Les objets récemment alloués sont conservés dans la pouponnière puis, lorsque ceux-ci sont retenus plus longtemps, ils sont déplacés vers une génération plus ancienne, et ainsi de suite. Seule la pouponnière est parcourue lors de chaque récupération et lorsque celle-ci est remplie, elle déclenche une récupération dans la génération au-dessus d'elle. Ainsi, cet algorithme est mieux adapté aux cas d'utilisations

les plus fréquents, car les objets récemment alloués ont souvent une courte durée de vie moyenne, contrairement aux objets déjà retenus, qui eux possèdent une meilleur chance d'être conservés à long terme.

Tout comme les deux techniques précédentes, les ramasse miettes générationnels nécessitent l'arrêt du programme durant l'opération récupération, qui peut, dans le pire des cas, s'effectuer sur toute la mémoire. Ainsi, les gestionnaires de mémoire générationnels améliorent les performances du temps de récupération moyen, au détriment du temps de récupération de toute la mémoire (pire cas).

3.6.1.4 Gestionnaires incrémentaux

Afin de remédier au problème d'arrêt complet de l'exécution du programme, un nouveau type de ramasse miettes a été développé : des gestionnaires de mémoire incrémentaux ou temps réels [35]. Ceux-ci peuvent utiliser un algorithme générique de marquage à trois couleurs de la mémoire [36] afin d'effectuer partiellement le travail de récupération et de poursuivre l'application. Cet algorithme fonctionne en subdivisant la mémoire en trois ensembles : l'ensemble blanc, l'ensemble gris et l'ensemble noir. Initialement, toutes les zones sont blanches, à l'exception des racines qui sont marquées comme grises. Ensuite, pour chaque élément de l'ensemble gris, l'algorithme marque chacune des zones accessibles par ceux-ci comme gris et marque finalement l'élément gris comme noir. À la fin, les zones mémoires toujours blanches sont inaccessibles et donc récupérées. Le travail de marquage peut être fait en étant subdivisé de manière à traverser un nombre fini de noeuds. Il en résulte donc que les pauses dues à la récupération peuvent être bornées supérieurement.

Afin d'entremêler le travail du récupérateur de mémoire et du programme exécuté (nommé le mutateur), certaines précautions doivent être prises. Pour ce faire, on doit éviter que le mutateur modifie des pointeurs en mémoire tel qu'il serait possible d'accéder à des espaces marqués comme étant non-accessibles et donc qui seront ultérieurement récupérés. Deux approches sont alors envisageables : des barrières de lecture ou d'écriture

sur les zones mémoires. Les barrières de lecture assurent que seuls des pointeurs noirs sont retournés. Ces dernières sont généralement très coûteuses car les accès mémoire constituent environ 15 % des instructions [37]. Par contre, celles-ci doivent être utilisées pour des récupérateurs incrémentaux copiants, tel un *Stop And Copy* incrémental. Les barrières en écriture possèdent plusieurs variantes qui effectuent des compromis entre l'efficacité algorithmique et l'efficacité de récupération.

3.6.2 Conclusion

Malgré que les techniques de récupération automatique de mémoire impliquent des coût supplémentaires en utilisation du processeur et de la mémoire, le gain en productivité pour le développement est tellement supérieur qu'ils sont devenus adoptés par presque tous les nouveaux langages de programmation. Le fait que l'exécution du programme puisse être arrêtée durant le moment de la récupération est potentiellement problématique pour l'écriture de jeux vidéo, car ceux-ci se doivent d'avoir des bonnes performances et ce, de manière continue. Les algorithmes de gestion de mémoire incrémentaux améliorent le pire cas du temps de récupération en augmentant le coût moyen. Ainsi, les coûts apportés à l'exécution du programme doivent être considérés. Donc, l'utilisation de récupérateurs générationnels semble un bon choix pour la programmation de jeux vidéo car ils offrent de bonnes performances pour les utilisations normales de la mémoire, comme c'est généralement le cas pour des jeux vidéo.

3.7 Dynamisme du langage

Le langage Scheme est un langage à typage dynamique, faisant ainsi contraste aux langages fonctionnels de la famille *ML*. Un typage dynamique se distingue par des vérifications de types effectuées durant l'exécution d'un programme. Un typage statique, quant à lui, effectuera les vérifications de types durant la compilation. Ainsi, une importante différence entre les deux approches est le moment où les erreurs de programmation apparaissent. D'autre part, dû à la nature beaucoup plus stricte du typage

statique, la résolution d’erreurs est beaucoup plus difficile, car elle dépend uniquement d’informations connues statiquement.

Un typage statique trouve des erreurs de types beaucoup plus tôt dans le processus de développement et permet d’exécuter un code libre de toutes vérifications de types. Une particularité des systèmes typés statiquement consiste à avoir des structures de code plus rigides, demandant possiblement beaucoup d’efforts pour être modifiées [16].

Pour un typage dynamique, les erreurs ne surviendront que durant l’exécution du programme et ne pourraient jamais ou difficilement se produire. De plus, un typage dynamique implique des pertes de performance puisque les vérifications de types sont effectuées pendant l’exécution du programme. Par contre, des erreurs apparaissant durant l’exécution peuvent être beaucoup plus faciles à corriger si l’implantation est munie d’un bon système de débogage. Aussi, un typage dynamique rend plus accessible la possibilité d’interpréter le langage en plus de le compiler. L’interprétation permet un développement rapide et facilite le prototypage grâce à la possibilité de faire de la programmation en direct (*live coding*).

3.7.1 Interprétation et débogage efficace

Un langage typé dynamiquement a la possibilité d’être interprété puisque les types ne doivent pas nécessairement être corrects avant l’exécution. On peut ainsi évaluer les instructions les unes à la suite des autres dans un environnement global d’interprétation (*top level*). Cette particularité infuse beaucoup de puissance au langage. En effet, cela permet de tester les modules écrits de manière indépendante et interactive. Les cas extrêmes peuvent ainsi être vérifiés et il est donc possible d’avoir une bonne assurance sur le fonctionnement correct des algorithmes utilisés.

En cas d’erreurs, le système est conçu afin de les détecter et les diagnostiquer. Puisque l’environnement est dynamique, une grande puissance d’introspection et d’analyse est disponible. Le système Gambit-C permet non seulement d’inspecter la pile de continuations courantes lors d’une exception, mais permet aussi l’introspection d’envi-

ronnements de fermetures et la possibilité de rétablir ces environnements et d'exécuter du code arbitraire dans ceux-ci dans le but de trouver la source de l'erreur.

De plus, le dynamisme du langage peut pousser la puissance de récupération d'erreurs encore plus loin en combinant le système de débogage de Gambit-C avec un protocole établi sur *TCP-IP* permettant de faire du débogage à distance. On peut ainsi analyser une erreur s'étant produite sur un processus distant, qui pourrait potentiellement être exécuté sur un système embarqué ou un appareil ne disposant pas d'interface permettant la récupération d'erreurs. Par exemple, ce système a été utilisé afin de débogger des processus Scheme exécutés sur un téléphone cellulaire.

Il en résulte donc que malgré un faible coût en performance lors de l'exécution d'un programme, un système utilisant le typage dynamique apporte beaucoup de flexibilité aux programmeurs et leur fournit des outils efficaces pour le développement de logiciels.

3.8 Extensions utiles

Cette section présente un recueil des extensions du langage Scheme qui sont utilisées dans les chapitres suivants de ce mémoire. Ces dernières ne constituent pas nécessairement un atout précis quant à la programmation de jeux vidéo, mais apporteront plutôt une meilleure compréhension des exemples de code fournis.

3.8.1 Fermetures

Le système Gambit-C permet la définition de paramètres avec l'utilisation de mots clés. La figure 3.22 illustre un exemple de définition d'une telle fonction avec quelques appels à celle-ci. On constate qu'il est alors possible de définir des valeurs optionnelles qui possèdent une valeur par défaut.

La syntaxe des fonctions a aussi été étendue de manière à permettre la spécification d'un nombre de paramètres arbitraires. Pour ce faire, le mot clé `#!rest` ou encore le point (`.`) est utilisé pour délimiter les paramètres normaux du paramètre *reste*,

```

(define (f x y #!key z (sum (+ x y)))
  (list x y z sum))

(f 1 2)                => (1 2 #f 3)
(f 1 2 z: 'allo)       => (1 2 allo 3)
(f 1 2 z: 'allo sum: 'salut) => (1 2 allo salut)

```

FIGURE 3.22 – Utilisation du passage de paramètres par mots clés

qui contiendra une liste de tous les paramètres supplémentaires utilisés. La figure 3.23 illustre un exemple d'utilisation du paramètre *reste*.

```

(define (g x y . z) (list x y z))

(g 1 2)                => (1 2 ())
(g 1 2 3 4 5)          => (1 2 (3 4 5))

```

FIGURE 3.23 – Exemple d'utilisation du paramètre *reste*

Lorsqu'une fonction reçoit plusieurs arguments, il est possible d'utiliser la procédure `apply` afin d'appliquer la fonction en question à une *liste* d'arguments. Cette fonction attend une fonction comme premier argument et un nombre arbitraire d'arguments par la suite. Le dernier argument passé doit être une liste. Le contenu de cette liste sera inséré dans un appel à la fonction spécifiée. La figure 3.24 illustre quelques utilisations de la procédure `apply`.

```

(+ 1 2 3 4 5)                => 15
(apply + 1 2 '(3 4 5))       => 15
(apply + (map (flip + 1) '(1 2 3 4 5))) => 20

```

FIGURE 3.24 – Exemples d'utilisations de la fonction `apply`

3.8.2 Tables de hachage

Le système Gambit-C supporte aussi l'utilisation de tables de hachage. Ces dernières peuvent être créées par la fonction `make-table` et des valeurs peuvent y être insérées ou inspectées avec les fonctions `(table-set! <table> <key> <value>)` et `(table-ref <table> <key> [<default-value>])`. Il est aussi possible de spécialiser ces tables

afin qu'elles soient plus efficaces pour des clés étant des symboles ou des chaînes de caractères.

3.8.3 Tests d'égalité et opérations conditionnelles

L'égalité en Scheme entre deux objets appartenant aux types de base du langage peut toujours être vérifiée par la fonction `equal?`. Cette dernière supporte ainsi la vérification d'égalité sur plusieurs types d'objets et se trouve ainsi peu efficace. La fonction `eq?` permet la vérification d'égalité sur des pointeurs et puisqu'en Scheme un symbole ou un mot clé ne possède qu'une instance unique, ce prédicat peut être également utilisé pour distinguer ces derniers types.

L'opération conditionnelle de base en Scheme est effectuée par la forme spéciale `(if <cond> <true> [<false>])`. Plusieurs formes spéciales dérivent de cette dernière afin d'exprimer plus facilement certains cas typiques. Par exemple, lorsque plusieurs conditions sont imbriquées les unes dans les autres, la forme spéciale `(cond (<pred> <body>) ... [(else <body>)])` peut être utilisée. Cette dernière permet aussi de récupérer la valeur retournée par un prédicat s'évaluant à vrai en ajoutant `=>` tel qu'illustré dans la figure 3.25.

```
(let ((t (make-table test: eq?)))
  (table-set! t (list 'a) 'a)
  (table-set! t 'b 'bonjour)
  (cond ((table-ref t (list 'a) #f) 'allo)
        ((table-ref t 'b #f) => (lambda (msg) msg))
        (else 'bonsoir)))

=> bonjour
```

FIGURE 3.25 – Exemple d'utilisation d'une table de hachage spécialisée pour des clés symboliques et d'utilisation de la forme spéciale `cond`

3.8.4 Déclaration et utilisation d'objets

Le système Gambit-C apporte la définition d'objets implantant des fonctionnalités minimales d'un système de programmation orientée objet. La définition de telles structures se fait en utilisant la forme spéciale

```
(define-type <type-id> [<member-id>] ... [extender: <extender-name>])
```

Celle-ci effectue la création de fonctions d'accès et de modification de membres, de création d'instances et de vérification de types. L'utilisation du mot clé **extender:** permet aussi la création d'une nouvelle forme spéciale (au nom spécifié) qui permet la définition de sous-types héritant des membres de la classe en question, implantant ainsi un héritage simple. La figure 3.26 illustre l'utilisation de cette forme spéciale.

```
(define-type point x y extender: define-type-of-point)
(define-type-of-point circle r)

(define (add p1 p2) (make-point (+ (point-x p1) (point-x p2))
                                (+ (point-y p1) (point-y p2))))

(add (make-point 1 2) (make-circle 3 4 5)) => #< point #4 x: 1 y: 1>
(if (circle? (make-point 1 2)) 'ok 'no)    => no
```

FIGURE 3.26 – Exemple de création d'objets dans Gambit-C

3.8.5 Affichage de résultats

Plusieurs fonctions sont disponibles en Scheme afin d'effectuer l'affichage du résultat de l'évaluation d'une expression. Le standard Scheme fournit entre autres les fonctions **write** et **display** qui permettent un affichage bas niveau de données Scheme. Le système Gambit-C offre aussi les fonctions **pretty-print** (ou **pp**) afin de faire un affichage enjolivé d'une sortie de **write**. La fonction **println** permet l'affichage d'expressions composées. La figure 3.27 illustre des exemples d'utilisations de ces fonctions.

```

(write (list 1 'a "a"))      => (1 a "a")
(display (list 1 'a "a"))   => (1 a a)
(pretty-print (list 1 'a "a")) => (1 a "a")\n
(println "Bonjour " (+ 1 2) (map (flip + 1) '(1 2 3)))
=> Bonjour 3234

```

FIGURE 3.27 – Exemples d'utilisations de fonctions d'affichage

3.8.6 Interfaces aux fonctions étrangères

Une interfaces aux fonctions étrangères (*Foreign Function Interface* ou FFI) est disponible dans le système Gambit-C. D'une part, cette dernière permet l'importation et l'utilisation de fonctions C ou C++ présentes dans des bibliothèques. La figure 3.28 illustre une interface minimaliste à la bibliothèque C OpenGL et une utilisation de cette interface. On constate donc qu'une fois importées, les fonctions de cette dernière peuvent être utilisées comme n'importe quelle fonction Scheme.

```

(c-declare "#include <gl.h>")

(c-define-type GLenum unsigned-int)

(define GL_QUADS ((c-lambda () int "___result = GL_QUADS;")))

(define glVertex2d (c-lambda ( GLdouble GLdouble ) void "glVertex2d"))
(define glBegin (c-lambda ( GLenum ) void "glBegin"))
(define glEnd (c-lambda () void "glEnd"))

(glBegin GL_QUADS)
(glVertex2d 0 0) (glVertex2d 0 1)
(glVertex2d 1 0) (glVertex2d 1 1)
(glEnd)

```

FIGURE 3.28 – Exemple d'utilisation de l'interface à des fonctions étrangères de Gambit-C

D'autre part, il est aussi possible d'exporter des fonctions écrites en Scheme de manière à ce qu'elles soient utilisées à partir de code C ou C++.

3.8.7 Capture de continuations

Le système Gambit-C effectue une abstraction de la fonction `call/cc` avec la fonction `continuation-capture` [38]. Cette dernière retourne plutôt une structure de données contenant la continuation du calcul au moment de l'appel à celle-ci. Une valeur peut alors être retournée à la continuation obtenue en utilisant la fonction `(continuation-return <cont> <val>)`. De plus, en utilisant ce mécanisme, il est possible d'attacher une continuation à la suite d'un *thunk* en utilisant la fonction `(continuation-graft <cont> <thunk>)`. La figure 3.29 illustre une telle capture de continuation, ainsi que des exemples d'utilisations de cette dernière.

```
(define k #f)
(+ 1 (continuation-capture (lambda (cont) (set! k cont) 2)))

(continuation-return k 2) => 3

(let ((thunk (lambda () (display 'allo) (+ 2 3))))
  (+ 10 (continuation-graft k thunk))) => allo6
```

FIGURE 3.29 – Exemple d'utilisation de l'API améliorée de capture de continuations fournie par Gambit-C

3.9 Conclusion

Cet aperçu des particularités du langage fonctionnel Scheme démontre clairement que le langage répond grandement aux besoins établis par les jeux vidéo. Notamment, la puissance d'abstraction provenant de l'aspect de programmation fonctionnelle du langage permet d'obtenir une bonne modularité du code source. Aussi, son système de macros procédurales ajoute la possibilité de créer des langages spécifiques aux domaines répondant à des problèmes précis pour l'écriture de jeux vidéo. Le dynamisme du langage apporte des outils de développement puissants, comme un interprète et un débogueur à distance, mais aussi permet de faciliter le développement itératif avec une approche par prototypage. Le gestionnaire automatique de mémoire évite la présence

potentielle de beaucoup d'erreurs reliées à la gestion de mémoire, mais pourrait causer des problèmes de ralentissement lors de l'exécution de jeux. Finalement, la possibilité de réifier les continuations d'un programme en cours d'exécution offre beaucoup de puissance expressive au programmeur.

Le langage Scheme étant très simple, il est essentiel de créer des extensions de ce langage afin de bâtir un jeu vidéo sur une base solide pour répondre aux besoins spécifiques. L'expressivité du langage rend cette tâche facile. Ainsi, deux extensions au langage Scheme ont été développées dans le but de répondre spécifiquement aux besoins des jeux vidéo.

CHAPITRE 4

PROGRAMMATION ORIENTÉE OBJET

Scheme est un langage offrant aux programmeurs beaucoup de ressources à l'état brut et leur permet de se bâtir des outils personnalisés pour mieux répondre à leurs besoins.

Un paradigme qui n'est pas supporté directement par le langage Scheme est celui de la programmation orientée objet. Ce paradigme possède des structures de données hiérarchiques, les objets, favorisant la modularité et l'encapsulation de celles-ci. Cette modularité provient des abstractions résultant de la hiérarchie des méta-objets appelés classes. L'encapsulation permet de cacher les données qui devraient demeurer internes au module créé. Le sous-typage résultant de la création de classes permet entre autres l'utilisation générique de fonctions souvent appelées méthodes ou fonctions génériques.

Ces objets sont centrés sur la notion d'états qui évolue au fil de l'exécution du programme. Malgré le fait que la notion d'états d'un programme ne se marie pas bien avec la programmation fonctionnelle, l'utilisation d'objets pour représenter les entités dans un jeu est très naturelle et appropriée. En effet, ces entités possèdent des caractéristiques propres et qui évoluent dans le temps. Par exemple, on peut imaginer une balle qui possède une vitesse de déplacement et une position.

De plus, l'utilisation de méthodes permet, d'une part, de réutiliser des comportements communs à des objets d'une même hiérarchie de classes et, d'autre part, de choisir un comportement approprié pour un objet donné, via le *dispatch* de méthode qui sélectionne automatiquement la méthode la plus appropriée disponible pour un objet donné.

Le SRFI-9 [39] (*Scheme Request For Implementation*) suggère un système très rudimentaire d'objets qui fut implémenté et étendu par le système Gambit-C [9]. Ce système permet la construction de manière générique d'objets typés avec héritage simple. L'exemple de la figure 3.26 illustre l'utilisation de ce système. L'accès aux membres des objets issus de la même hiérarchie est ainsi possible et permet un poly-

morphisme minimal.

Ainsi, le système Gambit-C nous permet d'utiliser des concepts de la programmation orientée objet de manière rudimentaire. Cette approche ne fournit aucune possibilité d'attacher des méthodes aux objets qui sont au coeur du polymorphisme.

L'approche plus traditionnelle quant à la déclaration de méthodes de classes est celle utilisée par le langage de programmation Java [40]. Cette approche lie les méthodes à une classe et ses descendantes. Les appels de ces méthodes possèdent ainsi un argument caché (souvent nommé *self* ou *this*) qui est une instance de cette classe. Le *dispatch* dynamique s'effectue *uniquement sur le type de ce premier argument*. C'est l'approche appelée *dispatch* simple (*Single Dispatch*). Cette approche est bien adaptée pour une interaction centrée sur un seul objet à la fois, mais est très limitée pour des interactions génériques sur de multiples objets différents. L'exemple de la figure 4.1 illustre la manière d'utiliser des méthodes traditionnelles pour effectuer du *dispatch* multiple. On constate qu'une discrimination se fait implicitement sur l'instance de la classe associée à la méthode, mais que le *dispatch* sur d'autres objets doit être fait explicitement. L'exemple illustré est très simple puisqu'il s'agit de *dispatch* double sur deux types, mais pourrait devenir très complexe s'il s'agissait de faire une discrimination sur plus de deux objets.

```

class Toto {
    public static void main(String[] args){
        Titi t = new Titi(1); Blub b = new Blub(2);
        System.out.println("t.add(t) = " + t.add(t));
        System.out.println("t.add(b) = " + t.add(b));
    }
}
class Blub {
    int val;
    Blub(int x){ val = x; }
}
class Titi {
    int x;
    Titi (int x){ this.x = x; }
    int add(Object other){
        int y = -1;
        if (other instanceof Titi) { y = ((Titi)other).x; }
        else if (other instanceof Blub) { y = ((Blub)other).val; }
        return x+y;
    }
}

```

$t.add(t) = 2$
 $t.add(b) = 3$

FIGURE 4.1 – Exemple d'utilisation de méthodes à la manière traditionnelle (en Java)

Le langage Common Lisp [19] possède un système orienté objet, le Common Lisp Object System [41] (CLOS), qui diverge beaucoup de l'approche traditionnelle. Ce système offre une grande flexibilité de comportements et une riche introspection via un protocole de méta-objets [42] (*Meta Object Protocol*). L'aspect le plus intéressant de ce système est l'utilisation de fonctions génériques, éliminant l'utilisation de méthodes traditionnelles. Une fonction générique est un ensemble de fonctions possédant le même nom et ayant des spécifications optionnelles de types pour ses arguments. Lorsqu'un appel de fonction est fait sur une fonction générique, le système décide alors de l'instance la plus appropriée à utiliser dépendant des arguments actuels passés dans l'appel, du type de ces arguments et de leur nombre. On obtient donc un système très dynamique effectuant du *dispatch multiple*. La figure 4.2 présente la définition d'une fonction générique effectuant automatiquement le choix de la bonne instance en se basant dynamiquement

sur les types des arguments passés.

```
(defclass Titi () ((x :accessor x :initarg :x)))
(defclass Blub () ((val :accessor val :initarg :val)))

(defgeneric add (p1 p2))
(defmethod add ((p1 Titi) (p2 Titi)) (+ (x p1) (x p2)))
(defmethod add ((p1 Titi) (p2 Blub)) (+ (x p1) (val p2)))

(let ((p1 (make-instance 'Titi :x 1))
      (p2 (make-instance 'Blub :val 2)))
  (print (add p1 p1)) => 2
  (print (add p1 p2)) => 3)
```

FIGURE 4.2 – Exemple de *dispatch multiple* écrit en CLOS

Il est ainsi possible d'écrire de manière concise des fonctions génériques effectuant un *dispatch* sur plusieurs arguments à la fois. On peut aussi constater que les fonctions génériques ne sont pas définies dans l'espace de nom d'une classe en particulier. Ceci résulte en un meilleur découplage entre les classes et les opérateurs interagissant avec leurs instances.

La programmation orientée objet apporte ainsi beaucoup d'abstractions intéressantes. D'une part, les entités présentes dans un jeu vidéo sont aisément représentées par l'utilisation d'objets. En effet, les objets possédant des caractéristiques d'états se confondent parfaitement à la perception d'entités de jeux qui possèdent elles aussi des états variant en fonction de l'évolution du jeu. D'autre part, une programmation orientée objet permet de faire abstraction des comportements communs à plusieurs types d'entités présentes dans un jeu vidéo tout en gérant de manière générique leurs états. L'utilisation de fonctions génériques permettrait ainsi de modulariser ces comportements. Il serait donc très intéressant d'utiliser la programmation orientée objet afin d'écrire des jeux vidéo en Scheme de manière naturelle et modulaire.

Plusieurs systèmes objets sont déjà disponibles pour l'implantation de Scheme Gambit-C. La plupart sont inspirés de la puissance expressive du système CLOS. On retrouve entre autres le système orienté objet Meroon [43]. Ce système impose plusieurs limita-

Fonctionnalités désirées	<code>define-type</code>	Meroon	OOPS
Héritage multiple			✓
Constructeurs génériques			
Membres d'instances	✓	✓	✓
Membres de classe			
Fonctions génériques polymorphiques		✓	✓
<i>Dispatch</i> sur des valeurs			✓

TABLEAU 4.1 – Comparaison de la présence des fonctionnalités désirées par les implantations existantes de systèmes d'objets disponibles sur Gambit-C

tions dont le fait que la hiérarchie de classes est limitée à un héritage simple. Aussi, il n'est pas possible d'allouer des attributs de classes, communs à toutes les instances. OOPS [44] est un autre système objet disponible pour Gambit-C. Celui-ci offre plus de fonctionnalités que Meroon, dont la possibilité de créer des classes avec héritage multiple. Cette fonctionnalité est très pertinente pour l'écriture de logiciels, car elle permet l'agrégation de plusieurs modules de manière très élégante et offre beaucoup de flexibilité. Le tableau 4.1 présente une comparaison de ces systèmes objets sur la présence de fonctionnalités utiles pour le développement de jeux vidéo.

Le tableau 4.3 de la section 4.3 illustre les performances de ces deux systèmes. On constate que les systèmes *define-type* de Gambit-C et Meroon sont beaucoup plus performants que OOPS. Il en résulte donc qu'aucun de ces systèmes ne semble offrir de bonnes performances tout en donnant accès aux fonctionnalités de la programmation orientée objet de haut niveau énoncées dans le tableau 4.1. Ainsi, afin d'obtenir un compromis entre l'expressivité et les performances, un nouveau système d'objets a été développé. Ce dernier aura comme vocation de répondre aux besoins en efficacité d'un jeu vidéo, tout en apportant le plus d'outils et de puissance d'abstraction que possible. Ce dernier est décrit en détail dans le reste de ce chapitre.

4.1 Description du système

Le système de programmation orientée objet développé est fortement inspiré du Common Lisp Object System [41] et possède la plupart des propriétés intéressantes

de ce dernier. On retrouve entre autres de l'héritage multiple de classes, du polymorphisme d'instances de classes et surtout, des fonctions génériques à *dispatch* multiple personnalisé. Malgré ces fonctionnalités de haut niveau, une attention a été apportée aux performances, tout spécialement à l'accès aux membres d'instances, puisque ces opérations sont fréquentes dans un jeu vidéo. En effet, le développement préliminaire de *Space Invaders* (chapitre 6) nous laissait déjà sous-entendre que l'accès aux membres d'instances est une opération très fréquente qui se doit d'être efficace afin de ne pas ralentir la cadence du jeu.

Le reste de cette section fournit un API (*Application Programming Interface*) au système, tout en motivant l'ajout des fonctionnalités clés de celui-ci et en fournissant des exemples d'utilisations.

4.1.1 Définition de classes

La définition de classes, faite via la forme spéciale *define-class*, est une version simplifiée de la forme `defclass` de CLOS. La syntaxe de `define-class` est présentée dans la figure 4.3.

```
(define-class <class-name> (<super1> ...) <member1> ...)
```

FIGURE 4.3 – Syntaxe de la forme spéciale `define-class`

Le méta-symbole `<class-name>` représente le nom que portera la classe définie. Suivant ce nom se trouve la liste de super classes, données par leurs noms respectifs. Par la suite, les membres de la classe sont donnés. Ces membres peuvent être un attribut d'instance (`slot: <slot-name>`), un attribut de classe (`class-slot: <slot-name>`) ou un constructeur d'instances (`constructor: <fun>`).

Cette forme spéciale est très similaire à `define-type` fournie par Gambit-C et fut conçue pour que les fonctions utilitaires générées soient compatibles avec celles créées par `define-type`. Ainsi, une conversion d'un système basé sur les objets de *define-type* est extrêmement simple. Il ne suffit qu'à remplacer la définition du type par une

définition de classe de notre système d'objets. Les figures 3.26 et 4.4 illustrent cette conversion. Aussi, le choix d'utiliser des fonctions compatibles à `define-type` rend la base du système d'objets triviale à utiliser par des nouveaux utilisateurs déjà familiers avec Gambit-C.

```
(define-class point () (slot: x) (slot: y))
(define-class circle (point) (slot: r))

(define (add p1 p2) (make-point (+ (point-x p1) (point-x p2))
                                (+ (point-y p1) (point-y p2))))

(add (make-point 1 2) (make-circle 3 4 5)) => #(<class-desc> 4 6)

(if (circle? (make-point 1 2)) 'ok 'no)    => no
```

FIGURE 4.4 – Utilisation simple du système objet compatible avec la forme `define-type`

Les fonctions générées par la définition de la classe `point` sont :

- `(make-point <x-val> <y-val>)` : Construit une nouvelle instance de la classe avec `<x-val>` et `<y-val>` comme valeurs initiales d'attributs.
- `(make-point-instance)` : Construit une nouvelle instance *non initialisée* de la classe `point`.
- `(point-x <instance>)` : Accesseur de l'attribut `x` de l'instance passée en argument. Une autre fonction est similairement générée pour l'attribut `y`.
- `(point-x-set! <instance> <new-x-val>)` : Fonction de mutation de l'attribut `x`. Une autre fonction est similairement générée pour l'attribut `y`.
- `(point? <any-value>)` : Prédicat de type de la classe `point`. Cette fonction retourne la valeur `#t` si la valeur passée est une instance de la classe `point` ou de n'importe quelle sous-classe de celle-ci. Sinon, cette fonction retourne `#f`.

En plus de définir des attributs d'instances, il est possible de définir des attributs de classe, communs à toutes les instances de celle-ci. Ainsi, les informations communes

à toutes les instances peuvent être factorisées, tout en conservant l’encapsulation de l’information dans l’objet. Il est important de noter qu’après la déclaration de la classe, tous les attributs de classe se retrouvent dans un état non-initialisé. Ils doivent donc être initialisés avant d’être utilisés. L’héritage d’attributs de classe fait en sorte que chaque sous-classe possède sa propre version de l’attribut.

Puisque ces attributs sont communs à toutes les instances, ils ne nécessitent pas ces dernières pour être accédés ou modifiés. Toutefois, ces fonctions d’accès ou de modification de valeur d’attributs de classes permettent de prendre optionnellement une instance afin d’utiliser ces fonctions de manière polymorphique. La figure 4.5 illustre ce processus.

```
(define-class toto () (class-slot: t))
(define-class blub (toto))

(toto-t-set! 10) (blub-t-set! 'allo)
(toto-t)          => 10
(blub-t)          => allo
(toto-t (make-toto)) => 10
(toto-t (make-blub)) => allo
(toto-t-set! (make-blub) 'salut)
(toto-t (make-blub)) => salut
```

FIGURE 4.5 – Exemple d’utilisation dynamique d’attributs de classe

En plus des créateurs d’instances à la **define-type**, le système objet permet de définir un ou plusieurs constructeurs pour la création d’instances simplifiées. Ces constructeurs sont plutôt des fonctions d’initialisation de nouvelles instances. Elles *doivent* avoir un premier argument correspondant à l’instance à initialiser et peuvent posséder un nombre arbitraire d’arguments supplémentaires. Ces constructeurs sont en fait utilisés pour créer une nouvelle instance de la fonction générique **init!**. Donc, il est également possible d’utiliser la discrimination de type faite par les fonctions génériques, comme décrit dans la section 4.1.2. L’accès aux constructeurs se fait par la forme spéciale **new** qui s’occupe de créer la nouvelle instance de l’objet et d’appeler la fonction générique **init!**. La figure 4.6 illustre l’utilisation de constructeurs génériques d’objets.


```

(define-class titi () (slot: i))

(define-class toto ()
  (slot: t)
  (constructor: (lambda (self t) (toto-t-set! self t)))
  (constructor: (lambda (self (t titi)) (toto-t-set! self (titi-i t)))))

(let ((obj1 (new toto 1))
      (obj2 (new toto (new titi 2))))
  (toto-t obj1) => 1
  (toto-t obj2) => 2)

```

FIGURE 4.6 – Utilisation générique de constructeurs

4.1.2 Définition de fonctions génériques

Toujours inspiré fortement du modèle éprouvé du système objet de Common Lisp [41], le système objet développé possède des fonctions génériques permettant de faire du *dispatch* multiple sur les arguments de celles-ci. Le choix d’adhérer à la philosophie de Lisp s’explique non seulement par le fait que nous souhaitons que notre système objet puisse offrir le plus de puissance expressive possible aux programmeurs de jeux vidéo. En effet, plusieurs problèmes se retrouvant dans le développement de jeux vidéo peuvent être exprimés de manière très concise avec des fonctions génériques. Par exemple, la résolution de collisions est un comportement qui dépend des deux entités impliquées et donc l’utilisation d’une fonction générique pour effectuer la résolution du travail à accomplir se fait de manière très naturelle.

La déclaration d’une nouvelle fonction générique se fait par le biais de la forme spéciale

```
(define-generic <gen-fun-name>)
```

qui ne prend que le nom de la fonction générique définie, puisque le nombre de ses arguments peut être variable. Les déclarations d’instances de fonctions génériques se font par l’entremise de la forme spéciale

```
(define-method (<gen-fun-name> <arg-desc1> ...) <body>)
```

qui possède une syntaxe similaire à une définition de fonction, mais où la syntaxe des arguments a été étendue. Une description d'arguments est généralement une liste de la forme (`<arg-name> <arg-type>`), où le premier élément est le nom de la variable et le deuxième le type de celle-ci. Le type est normalement le nom de la classe à laquelle l'instance peut appartenir (incluant le polymorphisme). Une description d'arguments peut aussi être un symbole unique qui sera le nom de la variable liée à la valeur passée en argument. Puisque dans un tel cas aucune information n'est donnée sur l'argument à recevoir, aucune discrimination d'instance ne sera faite en utilisant cet argument, outre le fait qu'il s'agit d'un argument de plus. Le type de l'argument sera implicitement `*` et donc, une telle déclaration est strictement équivalente à (`<arg-name> *`). La figure 4.7 présente un exemple de déclarations et d'utilisations d'instances d'une fonction générique effectuant des additions sur des objets s'étalant sur plusieurs types et ayant un nombre variable de paramètres.

```
(define-class length () (slot: 1))
(define-class point () (slot: x) (slot: y))
(define-class circle (point) (slot: radius))

(define-generic add)
(define-method (add x y) (+ x y))
(define-method (add x y z) (+ x y z))
(define-method (add (p point) (l length))
  (+ (point-x p) (length-l l)))
(define-method (add (p1 point) (p2 point))
  (new point (+ (point-x p1) (point-x p2))
             (+ (point-y p1) (point-y p2))))

(let ((p (new point 1 2)) (l (new length 5)))
  (add 10 11)      => 21
  (add 10 11 12)   => 33
  (point-x (add p p)) => 2
  (add p 1))       => 6
```

FIGURE 4.7 – Exemple de déclarations et d'utilisations d'une fonction générique

Le type d'un argument de méthode peut être membre d'une syntaxe étendue permettant de faire de la discrimination d'instances de fonctions génériques sur la *valeur*

des paramètres passés. Cette syntaxe est l'une parmi :

- (**match-value:** <value>) : La valeur du paramètre actuel doit être égale (selon `equal?`) à <value> pour discriminer ce paramètre d'une fonction générique. L'expression <value> est considérée constante.
- (**match-member-value:** <class> <slot-name> <value>) : Le paramètre actuel doit être une sous-classe de <class> et la valeur de son attribut <slot-name> doit être égale (selon `equal?`) à <value>. L'expression <value> est considérée constante.
- (**or** <match1> ...) : La valeur du paramètre doit correspondre à *au moins une* des clauses <match> pour discriminer l'instance de la fonction générique correspondante. Les clauses <match> doivent être une déclaration de type correcte. On peut retrouver ainsi un nom de classe ou l'une de ces quatre clauses spéciales.
- (**and** <match1> ...) : La valeur du paramètre doit correspondre à *toutes* les clauses <match> pour discriminer l'instance de la fonction générique correspondante. Les clauses <match> doivent être une déclaration de type correcte. On peut retrouver ainsi un nom de classe ou l'une de ces quatre clauses spéciales.

Un exemple simple de discrimination par valeur de paramètres est donné dans la figure 4.8. Il est important de noter que la discrimination fonctionne bien lorsque les instances effectuent de la discrimination orthogonale, *i.e.* que les possibilités ne se recoupent pas entre elles. Une discrimination arbitraire pourrait se produire si plusieurs instances peuvent être discriminées de manière équivalente pour un même appel.

Une note importante liée à l'utilisation des fonctions génériques est que les instances discriminant sur une classe données doivent apparaître *après* la déclaration de celle-ci, sinon le comportement du système est indéterminé.

Puisque le système objet supporte de l'héritage multiple, il est possible qu'une fonction générique possède uniquement deux instances correspondant à des super classes

60

```
(define-class point () (slot: x) (slot: y) (class-slot: zero))
(point-zero-set! (new point 0 0))

(define-method (div (p point) (n (match-value: 0)))
  'division-by-zero!)

(define-method (div (p point) n)
  (new point (/ (point-x p) n) (/ (point-y p) n)))

(define-method (div (p (and (match-member: point x 0)
                           (match-member: point y 0)))
                  n)
  (point-zero))

(point-x (div (new point 1 2) 2))      => 1/2
(div (new point 1 2) 0)                 => division-by-zero!
(eq? (div (new point 0 0) 5) (point-zero)) => #t
```

FIGURE 4.8 – Exemple de discrimination d’instances de fonctions génériques par la valeur des arguments

d’une classe donnée, comme c’est le cas dans la figure 4.9. Les classes `colored` et `circle` sont toutes deux des super classes de `colored-circle` et il n’existe pas d’instance de la fonction générique `test` discriminant directement sur le type `colored-circle`. Ainsi, le système doit choisir entre les deux instances disponibles puisque ces dernières sont utilisables pour une instance de la classe `colored-circle`. L’algorithme discrimine alors comme suit : il considère que l’instance de la fonction générique qui possède *la plus profonde hiérarchie de classes* est la plus spécifique. Cette profondeur est approximée par la somme du nombre de super classes que possède chacun des arguments de l’instance de la fonction générique en question. Ainsi, l’instance `(test (x colored))` obtient une valeur de 0, puisque `colored` ne possède aucune super classe et l’instance `(test (x circle))` obtient une valeur de 1, puisque la classe `circle` est une sous classe de `point`. C’est donc l’instance discriminant sur un objet de type `circle` qui est choisie. Si les valeurs de discrimination sont égales pour plusieurs instances, alors c’est la méthode définie en dernier qui sera utilisée.

Finalement, il est possible d’utiliser un *cast* pour effectuer l’appel spécifiquement

```

(define-class point () (slot: x) (slot: y))
(define-class circle (point) (slot: radius))
(define-class colored () (slot: color))
(define-class colored-circle (circle colored))

(define-method (test (x colored)) 'colored)
(define-method (test (x circle)) 'circle)

(test (new colored-circle 1 2 3 'red)) => circle

```

FIGURE 4.9 – Exemple illustrant l’algorithme de sélection d’instances de fonctions génériques

désiré. La syntaxe d’utilisation est la suivante : (`<genfun> cast: (<arg1-ty> ...)` `<arg1> ...`). Le *cast* est passé comme mot clé dans l’appel de la fonction générique et son argument doit être une liste des types de l’instance à choisir. Un exemple d’utilisation du *cast* est illustré dans la figure 4.10. Les arguments de la fonction générique `init!` sont d’abord évalués. Il en résulte qu’une instance non initialisée de la classe `container` est créée ainsi qu’une instance de la classe `circle`. Le constructeur de la classe `circle` utilise un *cast* afin d’utiliser également le constructeur de la classe `point`. Ensuite, l’objet conteneur est initialisé en choisissant explicitement le constructeur attendant un argument de type `point`. L’objet correctement initialisé est finalement retourné.

```

(define-class point () (slot: x) (slot: y)
  (constructor: (lambda (self x y)
    (pp 'creating-point)
    (point-x-set! self x) (point-y-set! self y))))

(define-class circle (point) (slot: radius)
  (constructor: (lambda (self x y r)
    (pp 'creating-circle)
    (circle-radius-set! self r)
    (init! cast: '(point * *) self x y))))

(define-class container () (slot: obj)
  (constructor: (lambda (self (p point))
    (pp 'containing-a-point)
    (container-obj-set! self p)))
  (constructor: (lambda (self (c circle))
    (pp 'containing-a-circle)
    (container-obj-set! self c))))

(init! cast: '(container point)
  (make-container-instance)
  (new circle 1 2 3))
=> creating-circle
    creating-point
    containing-a-point
    <container-instance>

```

FIGURE 4.10 – Exemple d'utilisation de *cast* pour l'appel de fonction générique

4.1.3 Fonctions et formes spéciales utilitaires

Plusieurs fonctions et formes spéciales utilitaires sont disponibles afin de simplifier les tâches répétitives ou de permettre une introspection limitée avec le système objet. On retrouve plusieurs fonctions permettant d'obtenir de l'information sur les types d'instances ou de classes. Parmi celles-ci, les principales sont :

1. `(instance-of? <any-value> <class-name>)` : Prédicat similaire au mot clé de même nom en Java qui détermine si la valeur passée est une instance directe de la classe spécifiée. Cette fonction est beaucoup plus efficace que sa contrepartie `is-subclass?`.

2. (`is-subclass? <any-value> <class-name>`) : Prédicat similaire au précédent, qui retourne vrai si et seulement si la valeur est une instance de la classe spécifiée (incluant les instance polymorphes).
3. (`find-class? <class-name>`) : Retourne le descripteur de la classe spécifiée si elle existe, ou dans le cas contraire `#f`.
4. (`get-class-id <any-value>`) : Retourne le nom de la classe associée à la valeur spécifiée. S'il s'agit d'une instance de classe, alors le nom de la classe sera retourné, sinon le type universel `*` sera retourné.
5. (`get-supers <class-name>`) : Retourne la liste de toutes les super classes associées à la classe spécifiée.

Ainsi, l'introspection demeure limitée en comparaison à celle disponible par le protocole de méta-objet de CLOS, mais elle demeure tout de même fonctionnelle. Une meilleure introspection permettrait de développer des classes encore plus génériques, mais cette fonctionnalité n'est pas nécessaire pour le développement de jeux.

D'autre part, trois formes spéciales sont fournies pour simplifier les tâches répétitives pour la gestion d'instances de classes. Celles-ci sont :

- (`new <class-name> <value1> ...`) : Forme spéciale créant une nouvelle instance non-initialisée et appelant la fonction générique d'initialisation `init!` avec cette nouvelle instance et les valeurs actuelles spécifiées.
- (`update! <instance> <class-name> <slot-name> <fun>`) : Cette forme spéciale permet de modifier la valeur d'un attribut en fonction de sa valeur précédente. La fonction passée doit prendre un seul argument en paramètre, qui est la valeur courante de cet attribut. La valeur retournée par la fonction sera alors utilisée comme nouvelle valeur de celui-ci.
- (`set-fields! <instance> <class-name> ((<slot-name> <value>) ...)`) : Cette forme spéciale permet de modifier de manière plus concise la valeur de plu-

sieurs attributs en même temps. Elle est utile dans l'écriture de constructeurs d'instances.

La figure 4.11 illustre un exemple d'utilisation de ces formes spéciales.

```
(define-class point () (slot: x) (slot: y))

(define-class colored-circle (point) (slot: radius) (slot: color)
  (constructor:
    (lambda (self x y r c)
      (init! cast: '(point * *) self x y)
      (set-fields! self colored-circle ((radius r) (color c))))))

(let ((p (new colored-circle 1 2 3 'red)))
  (update! p point x (flip + 1))
  (point-x p)) => 2
```

FIGURE 4.11 – Exemple d'utilisation des formes spéciales fournies par le système objet

4.2 Implantation

Afin d'implanter ce système objet comme une extension directe du langage Scheme, les macros procédurales ont dû être utilisées. Il en résulte donc l'ajout d'un langage spécifique au domaine de la programmation orientée objet.

Puisque l'expansion de macros Scheme se fait lors de la compilation, il devient avantageux de faire le plus de travail possible durant cette expansion afin d'obtenir de bonnes performances lors de l'exécution. Par contre, afin d'effectuer beaucoup de travail lors de cette expansion, une grande quantité d'informations doivent être disponibles de manière statique, ce qui brime la philosophie dynamique du système objet. Ainsi, nous avons opté pour un compromis entre la division du travail effectué lors de la compilation et de celui fait lors de l'exécution du système objet.

De plus, les informations recueillies lors de l'expansion macro sont souvent nécessaires au fonctionnement du système lors de son exécution. Un passage d'informations entre ces deux mondes doit donc être fait. Afin d'y arriver, les structures correspondantes (descripteurs de classes, descripteurs de fonctions génériques, *etc.*) sont recréées durant

l'exécution. Pour distinguer plus facilement ces deux types de structures, les préfixes **mt** (*macro expansion time*) et **rt** (*runtime*) sont utilisés dans le code source.

Les sections qui suivent expliquent les grandes lignes de l'implantation de ce système objet et justifient les choix d'implantations qui ont dû être faits.

4.2.1 Implantation de **define-class**

Le fonctionnement de **define-class** est intimement lié aux structures de données utilisées pour conserver l'information sur les classes afin de transmettre celles-ci aux instances. L'expansion de la forme spéciale **define-class** résulte en la définition de plusieurs fonctions de création et d'accès aux données d'instances, mais aussi en la création d'un *descripteur de classe*. Ce dernier a pour rôle de conserver non seulement les super classes associées à la classe définie, mais aussi de conserver l'information permettant d'accéder aux attributs d'une instance. Les descripteurs de classes utilisés durant l'exécution sont une version légèrement simplifiée des descripteurs utilisés durant l'expansion de macros.

Le polymorphisme d'accès aux attributs d'instances est possible grâce à une indirection à partir de ce descripteur de classe. Malgré le fait qu'une telle indirection ralentit légèrement l'accès aux attributs, cette dernière permet aussi d'avoir des instances qui ne possèdent pas d'espaces inutilisés. Ainsi, des indices donnant l'emplacement des attributs dans les instances sont conservés dans les descripteurs de classes. Puisque l'héritage multiple est permis, ces indices peuvent varier d'une sous-classe à l'autre. Il en résulte donc que ces index d'indirection doivent toujours se retrouver au même endroit dans les descripteurs de classes afin que les accesseurs des super classes sachent où trouver cette information. La figure 4.12 contient un exemple de définitions de classes permettant d'illustrer les structures utilisées lors de l'exécution du système. La figure 4.13 schématise ces structures.

Il aurait été possible de faire en sorte que ces index soient toujours au même endroit

```

(define-class point () (slot: x) (slot: y) (class-slot: id))
(define-class length () (slot: l))
(define-class point3d (point) (slot: z))

(point-zero-set! 'toto)
(point3d-zero-set! 'titi)

(define p (new point 11 22)) => #(#(point () 1 2 toto) 11 22)

(define distance (new length 33))
=> #(#(length () unknown-slot unknown-slot unknown-slot 1) 33)

(define p2 (new point3d 44 55 66))
=> #(#(point3d (point) 1 2 titi unknown-slot 3) 44 55 66)

```

FIGURE 4.12 – Exemple concret de structure d’instance, descripteur de classe et de fonction d’accès à un attribut par indirection

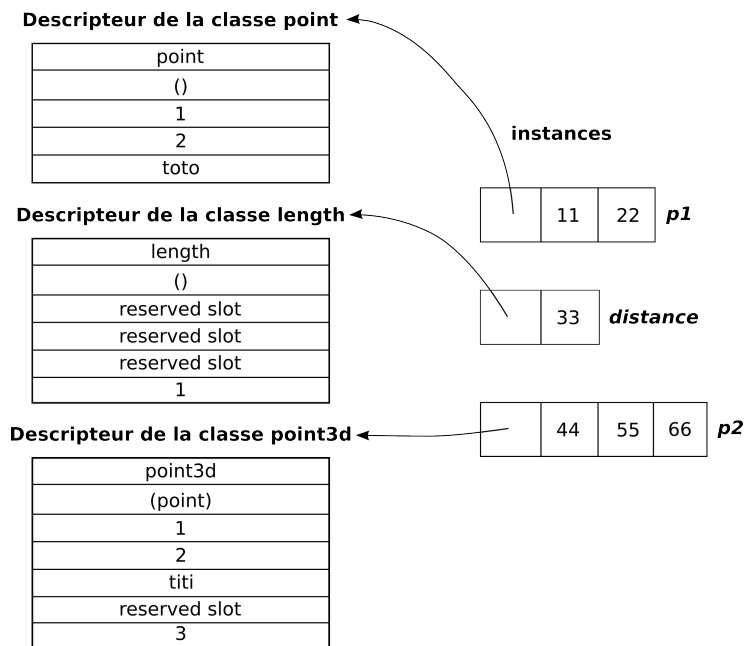


FIGURE 4.13 – Schématisation des structures créées dans la figure 4.12

pour une hiérarchie de classes donnée, mais nous avons simplifié le problème dans notre implantation. En effet, les index d’indirection se retrouvent au même endroit dans le descripteur de classe, peu importe la classe. La conséquence de cette simplification est que la taille des nouveaux descripteurs de classe augmente linéairement en fonction du nombre d’attributs qui ont été déclarés pour les classes précédemment définies. Une autre possibilité aurait été de toujours placer les attributs au même endroit dans les

instances. Puisqu'en général peu de classes sont utilisées en comparaison à la quantité d'instances créées, le choix effectué est un bon compromis entre l'utilisation de la mémoire et la rapidité d'accès aux attributs. Ces structures de données sont implantées en utilisant des vecteurs Scheme. Ceux-ci ont été choisis dans le but de donner un accès rapide aux attributs.

Lorsque la forme spéciale `define-class` est expansée, plusieurs informations sont extraites et traitées afin de générer correctement les fonctions de création et d'accès aux attributs. Entre autres, il faut obtenir la liste complète de tous les attributs hérités par les classes parents à la classe définie. Cette information est stockée dans une table de hachage globale à l'expansion macro nommée `mt-class-table` qui conserve l'information disponible pour chaque classe déjà créée. Cette information, stockée sous forme de méta-descripteurs de classe, n'est disponible que pour l'expansion macro. Ces descripteurs gardent une trace du nom des classes créées, de leur hiérarchie et une liste de positions relatives aux descripteurs de classes présents durant l'exécution. Ces positions relatives indiquent où trouver les index d'indirection associés à leurs attributs.

Lorsque les nouveaux attributs (non-hérités) sont inclus dans un descripteur de classe de l'exécution, un compteur global (expansion macro) est utilisé pour obtenir la position du nouveau champ dans le vecteur correspondant à ce descripteur. Une fois toutes les positions déterminées, la génération du descripteur de classe de l'exécution peut être entamée. Un deuxième index, local celui-ci, est utilisé pour trouver l'indice d'indirection des attributs dans l'instance. Ainsi, l'index global pointe vers l'endroit où placer l'index d'indirection des attributs d'instance dans le descripteur de classe de l'exécution. Pour leurs parts, les index d'indirection indiquent où trouver les attributs dans les instances.

Une fois cette information en main, toutes les fonctions d'accès et de modification d'attributs peuvent être générées sans problème. La plupart contiendront des indirections pour l'accès aux champs, sauf pour les attributs de classe qui eux, se retrouvent directement dans les descripteurs de classe. La figure 4.14 donne un exemple concret

de code généré effectuant l'accès au champ `x` de la classe `point` définie dans la figure 4.12. Au total, trois indirections sont nécessaires à l'obtention de la valeur d'un attribut d'instance. La première indirection obtient le descripteur de classe de l'instance en question. La deuxième cherche dans ce descripteur l'index d'indirection qui donnera la position de l'attribut dans l'instance. Puisque la position de cet index d'indirection est globale, elle est directement insérée dans le code généré. Dans cet exemple, cette position se trouve à l'index 2 dans le descripteur de classe. Finalement, la valeur de l'attribut est obtenue en effectuant l'indirection sur l'instance.

```
(lambda (#:obj518)
  (vector-ref #:obj518
    (vector-ref (instance-class-descriptor #:obj518)
      2)))
```

FIGURE 4.14 – Code généré pour l'accès à l'attribut `x` de la classe `point` de la figure 4.12

Deux constructeurs d'instances sont aussi générés. Le premier (`make-<class-name>`) construit et initialise *tous* les champs d'une instance de la même manière que procède le constructeur d'instance pour la forme `define-type` de Gambit-C. Le deuxième (`make-<class-name>-instance`) ne prend aucun argument et produit une instance non-initialisée qui pourra être passée à `init!`, la fonction générique d'initialisation d'objets. Par défaut, une instance de cette fonction générique est créée, et fait exactement le même travail que le premier constructeur. De plus, une fonction très rudimentaire d'inspection d'instance de classe est implantée sous la forme d'instances de la fonction générique `describe`.

Les choix d'implantation qui viennent d'être énumérés ont été faits dans le but d'obtenir un compromis entre la rapidité du système et la puissance expressive auquel il donne accès. Par exemple, on constate qu'une indirection est faite sur une instance afin d'obtenir son descripteur de classe, tandis qu'un pointeur vers ce descripteur de classe aurait pu être placé directement (voir la figure 4.14). Par contre, si un pointeur s'était

retrouvé directement dans la fonction d'accès `point-x`, il n'aurait alors plus été possible de l'utiliser de manière polymorphique avec les sous-classes de la classe `point` puisqu'il n'est pas certain que l'attribut `x` se retrouvera au même endroit dans une instance de ces sous-classes. Puisqu'un accès aux champs d'un vecteur est une opération efficace en Scheme, nous avons préféré opter pour la présence du polymorphisme qui apporte beaucoup en terme d'expressivité.

Puisque le système objet est implanté en utilisant l'expandeur de macros `define-macro` de Gambit-C, les données conservées sur les classes définies durant cette expansion sont perdues lorsque celle-ci est terminée. Par conséquent, il n'est pas possible de modifier la hiérarchie de classes définie dans un module Scheme (un fichier) à l'intérieur d'un autre module. Par exemple, une classe `Alpha` définie dans un fichier `A` ne pourrait pas être sous-classée dans un autre fichier `B`. Toutefois, il est possible d'utiliser les hiérarchies de classes créées dans un premier module à l'intérieur d'un autre module. L'implantation du support intermodulaire du système objet n'a pas été abordée dû à la complexité de la tâche et à la potentielle baisse de performance qui pourrait découler d'un tel ajout.

4.2.2 Implantation de `define-generic`

De manière similaire à la définition de nouvelles classes, des informations sur les fonctions génériques définies et sur leurs instances sont conservées dans une table de hachage globale durant l'expansion macro et se nomme `mt-meth-table`. Ces informations seront par la suite transférées vers l'exécution du programme sous la forme de structures propres à chaque fonction générique qui se nomme `<genfun-name>-meth-table`. Ces structures contiennent le nom de la fonction générique, une table de hachage permettant de vérifier rapidement l'existence d'une instance et une liste *triée* des instances permettant d'accélérer le polymorphisme de fonctions génériques. La clé de cette table de hachage est basée sur le type des arguments de l'instance stockée.

L'expansion de la définition d'une nouvelle fonction générique résulte donc en la création de cette structure et d'une fonction portant le nom de la fonction générique,

qui a pour but de faire le choix de la bonne instance à partir des paramètres qui lui sont passés. La fonction de *dispatch* générée pour la fonction générique `init!` est illustrée dans la figure 4.15. Le code présenté a été légèrement modifié afin d'illustrer uniquement l'essence du *dispatch* effectué. On constate que dans un premier temps, on cherche dans la table de hachage de cette fonction générique si une instance associée aux types des arguments passés (ou du *cast* effectué) existe. Si ce n'est pas le cas, on cherche explicitement une instance admissible de manière polymorphe à ces arguments.

```
(lambda (#!key cast . args)
  (let ((types
        (if cast
            cast
            (cond ((= args-nb 0) '())
                  ((= args-nb 1) (list (get-class-id (car args))))
                  ...
                  (else (map get-class-id args))))))
    (cond
      ((or (generic-function-get-instance init!-meth-table types)
           (find-polymorphic-instance? init!-meth-table
                                         args types))
       => (lambda (method)
            (cond ((= args-nb 0) ((method-body method)))
                  ((= args-nb 1) ((method-body method) (car args)))
                  ...
                  (else (apply (method-body method) args)))))
      (else (error ...)))))
```

FIGURE 4.15 – Code expansé effectuant le *dispatch* dynamique pour la fonction générique `init!`

L'algorithme qui détermine quelle instance est la plus spécifique pour un ensemble d'argument donné est simple : on cherche à trouver la première instance dont chacun des types est considéré équivalent au type des arguments actuels *dans une liste triée* des instances en fonction de leur spécificité. Le critère de spécificité est déterminé par la profondeur dans la hiérarchie du type des arguments de l'instance, *i.e.* par la somme du nombre de classes parents pour chacun des types des arguments passés. En ce qui concerne les types spéciaux comme `match-value`, ces types sont considérés

comme étant très spécifiques et donc prioritaires à une simple correspondance du type d'un objet. La figure 4.16 illustre le code utilisé pour réaliser ce *dispatch* polymorphique.

```
(define (find-polymorphic-instance? genfun
                                     actual-params actual-types)
  (let ((args-nb (length actual-params))
        (sorted-instances
         (generic-function-sorted-instances genfun)))
    (exists (lambda (method)
              (and (= (length (method-types method)) args-nb)
                   (equivalent-types? (method-types method)
                                       actual-params
                                       actual-types))))
      sorted-instances)))
```

FIGURE 4.16 – Implantation de la recherche d'instances polymorphiques d'une fonction générique

À un certain moment la fonction `call-next-method` avait été implantée en utilisant le mécanisme de variables à portée dynamique de Gambit-C. Par contre, l'exécution du corps des méthodes à l'intérieur d'un environnement dynamique ralentissait l'exécution de celles-ci d'un facteur d'environ 15 %. Par conséquent, l'implantation de cette fonction a été abandonnée. En effet, celle-ci permet d'écrire des corps de méthodes légèrement plus génériques, mais l'utilisation du `cast` permet de faire sensiblement le même travail sans avoir à subir des pertes de performance.

4.2.3 Implantation de `define-method`

L'expansion de la macro `define-method` consiste en la création d'une structure de données qui contient le corps de l'instance définie ainsi que les types associés aux arguments attendus. Cette structure existera durant l'expansion et sera recrée lors de l'exécution du programme afin d'être appelée par la fonction de *dispatch*.

opération	complexité algorithmique
création d'instances directes	$O(1)$
création avec constructeurs	$O(p)$
accès aux membres	$O(1)$
modification de membres	$O(1)$
<i>dispatch</i> direct	$O(p)$
<i>dispatch</i> polymorphique	$O(mp)$

TABLEAU 4.2 – Complexité algorithmique des opérations du système objet effectuées durant l'exécution

4.3 Performances

Ce système de programmation orientée objet a été écrit dans le but d'obtenir un compromis entre l'accès à des concepts de haut niveau et de bonnes performances lors de l'exécution, notamment en ce qui concerne l'accès aux membres d'instances de classes.

Dans un premier temps, les performances du système sont analysées de manière théorique. Le tableau 4.2 donne la complexité algorithmique des opérations utilisées. La création d'instances directes, *i.e.* utilisant les fonctions `make-<class-name>`, s'effectue en temps constant puisqu'il ne s'agit que de la création d'un vecteur. Par contre, la création d'instances par des constructeurs, *i.e.* en utilisant `(new <class-name> ...)`, possède la complexité des appels de fonctions génériques directs. Par « appel direct », il est sous-entendu que le type des arguments de l'appel correspond aux types des arguments formels, et non de manière polymorphe. Ces appels possèdent une complexité algorithmique linéaire en fonction du nombre de paramètres que possède l'instance (appelons-le p). En effet, l'obtention d'une instance directe est faite par l'entremise d'une recherche dans une table de hachage utilisant `equal?` à titre de comparaison. Cette complexité augmente d'un ordre supplémentaire lorsqu'aucune instance directe n'existe et qu'une instance polymorphique est recherchée. En effet, une recherche linéaire est effectuée sur chacune des instances existantes (m), puis la compatibilité du type des arguments doit être effectuée (p). Il est intéressant de noter que malgré la complexité quadratique de cette opération, le nombre de paramètres (p) est normalement très bas. L'accès et la modification de membres de classe se fait toutefois de manière constante.

Opération	define-type	Meroon	OOPS	class
Création d'instances directes	0.005	0.215	1.423	0.007
Création avec constructeurs	ND	ND	ND	1.318
Accès aux membres	0.006	0.126	3.120	0.014
Modification de membres	0.008	0.117	3.961	0.011
<i>Dispatch</i> direct ($p = 2$)	0.185	0.641	19.210	0.897
<i>Dispatch</i> direct ($p = 5$)	0.411	1.008	22.820	1.997
<i>Dispatch</i> polymorphique ($p = 2$)	0.188	0.705	18.995	3.578
<i>Dispatch</i> polymorphique ($p = 5$)	0.452	1.200	24.052	10.125

TABLEAU 4.3 – Temps d'exécution (en secondes) de 1 000 000 itérations de tests comparatifs sur les opérations du système de programmation orientée objets développé avec d'autres systèmes objet existants

Une comparaison pratique de performance a été effectuée avec d'autres systèmes objets disponibles pour Gambit-C afin de mieux connaître les constantes cachées dans la complexité des opérations effectuées durant l'exécution. Il est aussi intéressant de comparer les performances du système par rapport aux autres afin de confirmer l'obtention du résultat désiré. Ces résultats sont présentés dans le tableau 4.3. Ces derniers correspondent aux temps nécessaires à 1000000 d'exécutions de chacune des opérations principales du système objets sur un processeur unique de fréquence 1,6 Go ayant 1 Go de mémoire centrale. Les systèmes objets Meroon [43], OOPS [44] ainsi que les objets obtenus avec l'utilisation de la forme spéciale `define-type` de Gambit-C ont été utilisés pour effectuer cette comparaison. Le tableau 4.4 présente ces résultats d'une manière relative au meilleur temps d'exécution du test pour chaque opération vérifiée. Les tests effectués sur des fonctions génériques utilisent quatre instances de celles-ci pour notre système objet ($m = 4$).

Ces résultats démontrent bien que les objectifs d'implantation ont été atteints. En effet, le système développé (`class`) offre de bonnes performances quant à la création d'instances et la manipulation des membres de celles-ci. Malgré le fait que notre système est environ un cinq fois plus lent qu'un appel de fonction dans Gambit-C pour effectuer le *dispatch* direct de fonctions génériques, *i.e.* où une instance de la fonction générique existe pour les types des arguments passés à celle-ci. Une dégradation d'un facteur de cinq est aussi observée pour effectuer le travail supplémentaire lors d'appels de fonctions

Opération	define-type	Meroon	OOPS	class
Création d'instances directes	1.0	43.0	284.6	1.4
Création avec constructeurs	ND	ND	ND	1.0
Accès aux membres	1.0	21.0	520.0	2.3
Modification de membres	1.0	14.6	495.1	1.4
<i>Dispatch</i> direct ($p = 2$)	1.0	3.5	103.8	4.8
<i>Dispatch</i> direct ($p = 5$)	1.0	2.5	55.5	4.9
<i>Dispatch</i> polymorphique ($p = 2$)	1.0	3.8	101.0	19.0
<i>Dispatch</i> polymorphique ($p = 5$)	1.0	2.7	53.2	22.4

TABLEAU 4.4 – Comparaison de performances relatives au système le plus rapide (1.0) pour chaque opération du système objet

génériques nécessitant du polymorphisme.

Ainsi, les performances obtenues se comparent très bien à Gambit-C lui-même pour la création d'objets et l'accès aux membres de ceux-ci. Aussi, la performance d'appels de fonctions génériques, quoi que légèrement plus lente, se compare aussi à Meroon qui est réputé pour être bien optimisé. Donc, tout en offrant plus de fonctionnalités que ses compétiteurs notre système offre des performances adéquates en fonction des critères fixés.

4.4 Conclusion

Ainsi un système objet complet a été développé dans le but d'étendre le langage Scheme et d'y inclure le paradigme de la programmation orientée objets. Ce système objet permet la déclaration de classes avec héritage multiple, polymorphisme et fonctions génériques effectuant du *dispatch* multiple. Ces choix de caractéristiques ont été faits dans le but de donner le plus de liberté aux programmeurs de jeux vidéo, tout en fournissant de bonnes performances, tant en temps sur processeur qu'en utilisation de la mémoire.

Dans ce but, les structures de données utilisées pour l'implantation d'instances de classes ne possèdent pas d'espaces inutilisés et ne requièrent que trois indirections vectorielles afin d'accéder aux champs de celles-ci. De même, les fonctions génériques utilisent des mécanismes de tri pré-calculé afin d'accélérer le *dispatch* d'instances de fonctions

génériques de manière polymorphe.

Il serait maintenant intéressant de modifier le système afin qu'il supporte un protocole de méta-objets. Un tel protocole donnerait accès à une introspection très développée et permettrait aux utilisateurs de modifier le comportement du système selon leurs besoins.

Aussi, une modification du système afin de le rendre compatible avec des classes de manière inter-modulaire serait une grande amélioration en ce qui concerne la modularité du code.

Par contre, une attention particulière devrait être portée aux coûts en performance que ces modifications pourraient impliquer afin de respecter la philosophie de base de ce système objet.

CHAPITRE 5

SYSTÈME DE COROUTINES

Les *threads* constituent un outil de programmation important permettant l'exécution parallèle de code, de façon matérielle ou logicielle. L'utilisation de *threads* dans un jeu vidéo peut être très pertinente et permettrait d'exprimer de manière concise des concepts clés. Par exemple, l'implantation de parties multi-joueurs où ces derniers jouent à tour de rôle est bien représentée par un modèle parallèle, puisqu'il s'agit de deux parties distinctes jouées en même temps.

Le langage Scheme tel que décrit par le standard [8] ne fournit pas de système de *threads*. Par contre, le document SRFI 18 [45] décrit un *API* de système de *threads* concurrents. Gambit-C [9] supporte cet *API* sous forme de *threads* verts correspondant à une concurrence logicielle et non matérielle. Malheureusement, le coût d'utilisation des mécanismes de synchronisations explicites est très élevé en temps de développement et en utilisation du processeur lors de l'exécution.

Par contre, un système de coroutines (appelé aussi *threads* coopératifs) permettrait d'éviter d'avoir à spécifier explicitement ces synchronisations entre entités. En effet, si le flot de contrôle est changé durant des moments opportuns connus du programmeur, aucune synchronisation supplémentaire n'est requise pour assurer la validité d'accès concurrents à des sections critiques. En fait, le problème disparaît complètement. Ceci rend donc très attrayant de tels systèmes pour le développement de jeux vidéo. Il permettrait d'exprimer de manière simple des changements de contextes dans le jeu, ou même de modulariser le comportement de chacune des entités du jeu. Afin de conserver la modularité du jeu, on peut imaginer une cascade de systèmes de coroutines dont un premier niveau permettrait l'implantation de parties multi-joueurs et un second niveau permettrait l'implantation du comportement d'entités dans le jeu. La figure 5.1 illustre de manière générique cette idée. Dans ce schéma, les rectangles noirs représentent des systèmes de coroutines et les coroutines sont représentées avec des lignes pointillées.

Ainsi, chaque joueur est lui-même un système de coroutine qui gère l'état de la partie de ce dernier. L'utilisation de systèmes imbriqués permettrait ainsi d'exprimer chaque concept de manière modulaire et complètement transparente.

Système de coroutines primordial

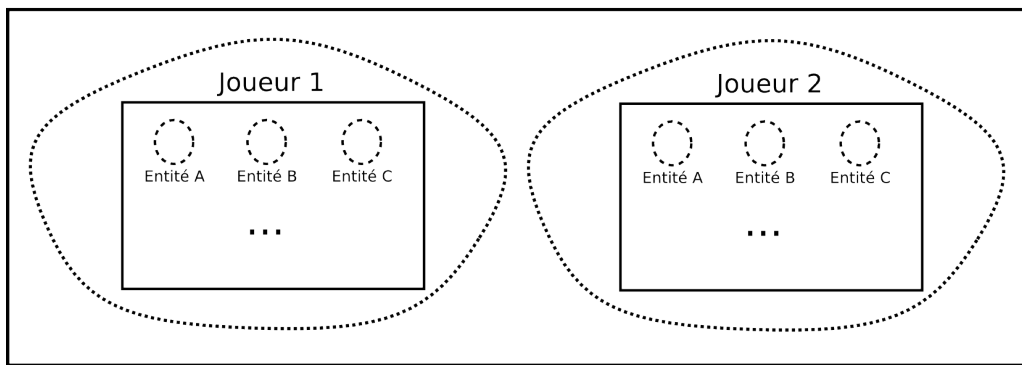


FIGURE 5.1 – Exemple générique d'utilisation de systèmes de coroutines cascades

Gambit-C offre la possibilité de transformer son système de *threads* en système de coroutines lorsque le délai de changement de contexte automatique est choisi comme étant l'infini. Toutefois, il est impossible d'utiliser plusieurs systèmes de coroutines avec cette approche. Ainsi, l'utilisation directe des fonctionnalités de Gambit-C limite celle de coroutines au développement d'un seul concept dans un jeu.

Afin de ne pas être limité par les fonctionnalités fournies par Gambit-C, l'expressivité du langage Scheme peut être mise à profit. En effet, grâce à la réification de continuations du calcul, la tâche de l'implantation d'un système de coroutines fait sur mesure est tout à fait accessible et réalisable sans changer le langage. Ce chapitre présente un système de *threads* coopératifs développé dans le but de bien répondre aux besoins de jeux vidéo quant à l'écriture de code s'exécutant en concurrence de manière sécuritaire tout en permettant l'utilisation récursive de ce système.

5.1 Description du langage

Un système de *threads* coopératifs implique que le changement de contexte d'exécution associé aux changements de *threads* doit être fait de manière explicite par les utilisateurs. Ainsi, ces changements de contextes peuvent être faits aux moments opportuns pour assurer l'intégrité des données. Toutefois, des synchronisations entre les différentes coroutines pourraient être encore nécessaires pour bien orchestrer l'exécution de ces dernières. Afin de permettre aux coroutines de communiquer entre elles, une synchronisation par passage de messages avec *pattern matching* similaire à celle utilisée dans les systèmes Termite [46] ou Erlang [47] a été adoptée. Ce mécanisme permet de synchroniser naturellement et de manière élégante les coroutines.

De plus, le système a été conçu pour être utilisé récursivement. Il est donc possible d'avoir une coroutine qui sera elle-même un système de coroutine, et ainsi de suite. Cette fonctionnalité a été implantée pour que le système soit le plus générique possible. De plus, l'idée de systèmes récursifs est aussi très près du langage Scheme dans lequel la récursion de fonctions est très courante et commune grâce à l'implantation de l'optimisation d'appels terminaux.

Un tel système de coroutines peut être également vu comme étant un système de simulations basées sur les agents, par opposition aux simulations basées sur des événements discrets. Le comportement des entités est alors représenté par chacune des coroutines du système. Dans cette optique, la notion de temps du système développé a été abstraite grâce à l'introduction de compteurs de temps (*timers*). Il est donc possible de choisir non seulement une granularité temporelle en spécifiant la fréquence de ce compteur, mais aussi de spécifier un facteur d'accélération permettant d'accélérer la simulation en cours.

Le système se résume à un ordonnanceur de coroutines qui utilise une file de coroutines prêtes à exécuter pour choisir la prochaine devant prendre le contrôle. Aussi, une file de coroutines en attente sur le temps et une sur des conditions sont disponibles. Lorsqu'une coroutine décide de passer la main à la suivante, ou lorsqu'elle ne doit plus

attendre après le temps ou une condition, elle se fait reléguer à la fin de la file d'attente de coroutines prêtes. La figure 5.2 illustre l'architecture globale du système.

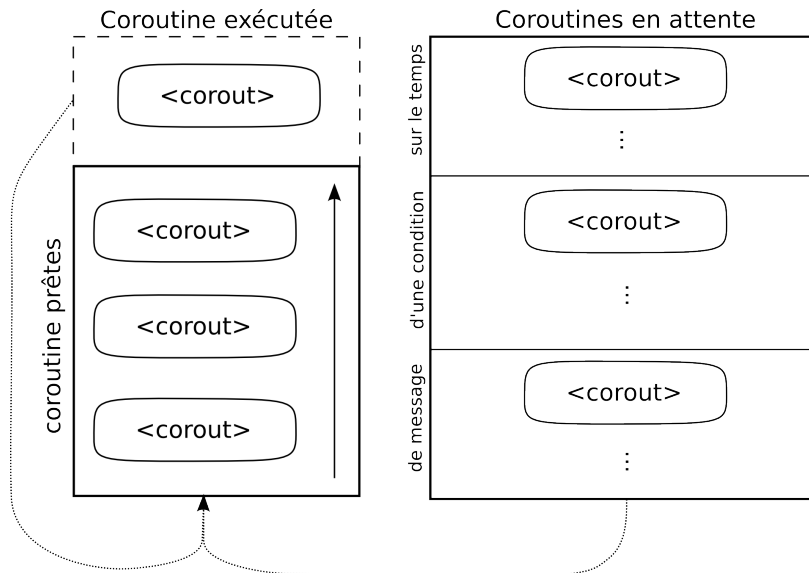


FIGURE 5.2 – Architecture globale du système de coroutines

Les sections suivantes décrivent un *API* permettant l'utilisation du système de coroutines développé.

5.1.1 Création de coroutines

Les coroutines sont des objets en eux-mêmes et peuvent être créées de manière externe au système de coroutines pour, par la suite, être ajoutées à celui-ci. L'avantage de cette approche, utilisée aussi par le système de *threads* de Gambit-C, réside dans le fait que les objets correspondant aux *threads* peuvent être initialisés à l'avance et conservés jusqu'au moment opportun de leurs ajouts au système. Une nouvelle instance de coroutine s'effectue avec

```
(new-corout <corout-id> <thunk>)
```

où `<corout-id>` est un symbole permettant d'identifier la coroutine et le dernier argument est un *thunk* qui contient le corps de l'exécution de cette coroutine.

La figure 5.3 illustre un diagramme des états possibles pour une coroutine avec les transitions correspondantes. Sur chacune des transitions, un exemple d'appel de fonction effectuant cette transition est donné. On constate ainsi qu'une coroutine est soit prête, en attente ou terminée. L'attente d'une coroutine peut être conditionnelle à la réception d'un message, au relâchement d'une sémaphore ou à un certain délai de temps prescrit. De plus, l'attente sur le temps peut être interrompue par la réception d'un message. Cet état se produit lorsqu'une attente de message est effectuée avec une attente bornée.

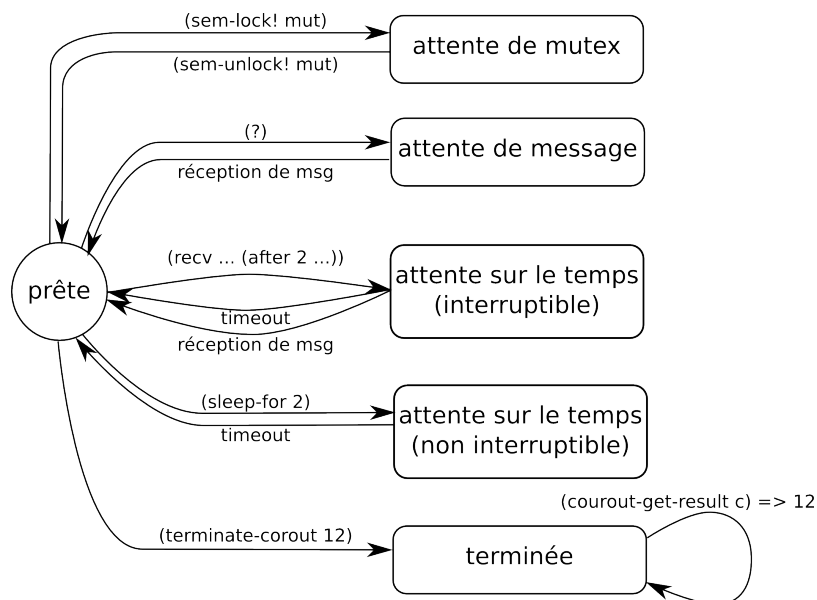


FIGURE 5.3 – Diagramme d'états d'une coroutine

L'objet créé pourra être par la suite ajouté à un nouveau système par la fonction d'initialisation `boot` ou encore par une coroutine d'une simulation existante avec la fonction `spawn-brother`. Ces fonctions sont décrites dans les sous-sections qui suivent.

5.1.2 Démarrage du système

Le système de coroutines peut être démarré en utilisant la fonction

```
(boot <list of corout> [<timer>])
```

dont le premier argument doit être une liste contenant les coroutines présentes lors du démarrage du système et, optionnellement un objet correspondant à un compteur de temps pour le système. L'ordre dans la liste est significatif car il indique le rang selon lequel seront enfilées les coroutines prêtes. Ainsi, la première coroutine de la liste sera la première à être exécutée. La création de *timers* se fait par un appel à la fonction

```
(start-timer! <period> [time-multiplier: <time-mult-value>])
```

qui s'occupe de créer un objet faisant abstraction du temps. Ce dernier sera rafraîchi selon la période (en secondes) indiquée en temps réel. On peut ainsi faire varier la granularité de la simulation afin de l'adapter aux besoins existants. Le paramètre optionnel `<time-mult-value>`, si spécifié, accélérera la simulation de ce facteur. Ainsi, le temps dans la simulation passera deux fois plus rapidement que le temps réel en utilisant une valeur de multiplicateur de temps de deux. La fonction `(pause-timer! <timer>)` peut être utilisée afin de mettre en pause et de repartir le compteur de temps.

La figure 5.4 présente un exemple de démarrage d'un système de coroutines où le temps sera augmenté toutes les secondes et où le résultat final de la simulation sera la somme des valeurs retournées par chacune des coroutines.

```
(let* ((c1 (new-corout 'c1 (lambda () (display 1))))
      (c2 (new-corout 'c2 (lambda () (display 2))))
      (c3 (new-corout 'c3 (lambda () (display 3)))))
  (boot (list c1 c2 c3))) => 123
```

FIGURE 5.4 – Exemple de démarrage d'un système de coroutines

Ce système a été conçu afin de permettre la création de systèmes de coroutines de manière cascades, *i.e.* qu'il est possible de créer des coroutines qui sont elles-mêmes des systèmes de coroutines. De tels systèmes récursifs comportent de nombreux avantages. En effet, l'utilisation de la récursivité facilite grandement la modularité. Il devient possible de modulariser les différents aspects d'utilisation dans leurs propres systèmes. Par exemple, dans le cadre de jeux simples, on peut implanter des parties multi-joueurs

en utilisant une coroutine par partie. Ces dernières peuvent alors elles-mêmes contenir un système de coroutines utilisé pour la partie.

La procédure de création de ces derniers est obtenue en démarrant un nouveau système de coroutines à l'intérieur d'une coroutine existante. Lorsque des coroutines démarrent un nouveau système de coroutines, l'objet d'abstraction du temps (*timer*) du système primordial est automatiquement utilisé pour les systèmes récursifs de manière à conserver une cohérence entre ceux-ci. Par contre, un problème se pose : il devient alors impossible de retourner le contrôle aux coroutines appartenant au système primordial, puisque (**yield**) ne fera que changer de contexte des coroutines du sous-système.

Ce problème est résolu par l'utilisation de la fonction (**super-yield**) qui effectue un changement de contexte pour le système de coroutines courant. Si le système courant ne se retrouve pas dans un système cascadié, alors rien ne se produira. Une fonction similaire, (**super-kill-all!** <return-value>) terminera *tous* les systèmes de coroutines se retrouvant dans l'arborescence de système de coroutines présentement actives. Un exemple de système de coroutines en cascade est donné dans la figure 5.5. On constate que les appels à **super-yield** ont bien effectué les changements de contextes des coroutines hôtes des sous-systèmes de coroutines (**s1** et **s2**).

```

(let* ((ret (lambda (x) (pp `(now returning: ,x)) x))
      (c1 (new-corout 'c1 (lambda () (ret 1))))
      (c2 (new-corout 'c2 (lambda () (super-yield) (ret 2))))
      (c3 (new-corout 'c3 (lambda () (super-yield) (ret 3))))

      (c4 (new-corout 'c4 (lambda () (ret 4))))
      (c5 (new-corout 'c5 (lambda () (super-yield) (ret 5))))
      (c6 (new-corout 'c6 (lambda () (ret 6))))

      (s1 (new-corout 's1 (lambda () (boot (list c1 c2 c3)))))
      (s2 (new-corout 's2 (lambda () (boot (list c4 c5 c6))))))
(boot (list s1 s2))
(fold-l + 0 (map corout-get-result (list c1 c2 c3 c4 c5 c6))))

=> (now returning: 1)
    (now returning: 4)
    (now returning: 2)
    (now returning: 5)
    (now returning: 6)
    (now returning: 3)
    21

```

FIGURE 5.5 – Exemple de démarrage de système de coroutines en cascade

5.1.3 Manipulation du flot de contrôle

Puisque le contrôle du flot d'exécution entre les coroutines doit être explicité par l'utilisateur, une bonne diversité de fonctions permettent la manipulation de celui-ci. Il est important de noter que, sauf avec avis contraire, toutes les fonctions décrites dans cette section doivent être exécutées par une coroutine, *i.e.* dans leur corps.

La fonction la plus simple de manipulation du flot d'exécution est (**yield**). Celle-ci arrête temporairement la coroutine actuelle et passe le contrôle à la *prochaine* coroutine disponible. Cette dernière est déterminée en utilisant la prochaine coroutine présente dans la file de coroutines prêtes, tel qu'illustré dans la figure 5.2. La coroutine effectuant l'appel à la fonction **yield** sera alors placée à la fin de cette file d'attente. Il en résulte que le retour du contrôle à cette coroutine ne sera exécuté uniquement que lorsque toutes les autres coroutines prêtes auront décidé de passer la main à la coroutine suivante. Il est également possible de choisir explicitement à quelle coroutine le contrôle sera passé avec la fonction (**yield-to <corout>**).

Une coroutine peut aussi décider de se mettre en attente pour une certaine période de temps avec la fonction (`sleep-for <sec>`) pour un certain nombre de secondes. Il en résultera que la coroutine actuelle sera placée dans une file d’attente séparée et ne sera réenfilée dans la file de coroutines prêtes uniquement que lorsque le délai prescrit sera dépassé. Cela n’implique pas que la coroutine sera exécutée à ce moment-là, car si cette file contient déjà des coroutines en attente, elle devra attendre son tour pour poursuivre son exécution. Puisqu’une telle attente peut varier en fonction de la plateforme utilisée, elle introduit un certain indéterminisme dans le système.

Il est également possible pour une coroutine d’ajouter de nouvelles coroutines dans le système en utilisant les fonctions :

```
(spawn-brother <corout>)
(spawn-brother-thunk <corout-id> <thunk>)
```

La première ajoute simplement la coroutine spécifiée dans la file de coroutines prêtes. Celle-ci *ne doit pas* déjà être présente dans le système. La deuxième fonction, `spawn-brother-thunk`, sert de sucre syntaxique en créant une nouvelle coroutine ayant comme identificateur `<corout-id>` et la fonction spécifiée comme corps. Par la suite, la coroutine nouvellement créée sera ajoutée au système.

Dans le cadre de l’écriture de jeux vidéo, la modularité des coroutines est très importante. Afin de faciliter l’écriture modulaire de corps des coroutines, des formes spéciales permettant d’utiliser un style CPS ont été ajoutées. Ces dernières permettent l’altération du fil d’exécution d’une coroutine en modifiant la continuation de son calcul par celui d’une autre coroutine ou d’un *thunk*. Ainsi, un appel à la forme spéciale (`continue-with <corout>`) fera en sorte que la continuation de la coroutine devienne celle de cette autre coroutine. De manière similaire, (`continue-with-thunk! <thunk>`) utilisera la fonction passée comme continuation du calcul. La figure 5.6 présente un exemple utilisant ces deux formes spéciales. Lorsque la coroutine `c1` s’exécute, sa continuation change pour celle de `c2` qui change à son tour pour la fonction `t`. Ainsi, la fin du corps des coroutines `c1` et `c2` ne sera jamais exécutée puisque leurs continuations

ont été altérées.

```
(let* ((t (lambda () (pp 'bonjour)))
      (c2 (new corout 'c2
                    (lambda () (continue-with-thunk! t) (pp 'allo ) 2)))
      (c1 (new corout 'c1
                    (lambda () (continue-with c2) (pp 'salut) 1))))
  (boot (list c1))) => bonjour
```

FIGURE 5.6 – Exemple de modification de la continuation d’une coroutine

Finalement, une coroutine peut changer sa continuation de manière plus drastique en forçant la terminaison du fil d’exécution de celle-ci grâce à la fonction (`terminate-corout` `<return-val>`). La valeur de retour de la coroutine spécifiée en paramètre est conservée dans la structure de données de la coroutine et peut être obtenue en faisant appel à la fonction (`corout-get-result` `<corout>`). Il est aussi possible de complètement terminer le système actuel de coroutines en faisant appel à (`kill-all!` `<return-val>`). La valeur de retour finale de l’exécution du système sera alors la valeur spécifiée.

5.1.4 Environnement dynamique

Une certaine introspection est très utile pour un programmeur, surtout pour des fins de débogage. Un environnement dynamique est disponible pour les coroutines lors de leurs exécutions. Cet environnement donne accès à de l’information sur le système de coroutines en cours d’exécution. Il comprend les paramètres :

- (`current-corout`) : retourne l’instance de la coroutine actuellement exécutée.
- (`timer`) : retourne l’objet d’abstraction du temps associé avec le système courant.

La valeur actuelle du temps peut être obtenue en appelant plutôt la fonction (`current-sim-time`).

5.1.5 Système de communication inter-coroutines

Que ce soit afin d’informer une partie du pointage d’un autre joueur ou d’un vaisseau spatial attendant le bon moment pour effectuer son déplacement, la synchronisation

des entités d'un jeu est primordiale. Le mécanisme principal de synchronisation des coroutines pour ce système est basé sur le passage de messages, tel que fait dans le système de programmation distribué Termit [46]. Cette approche semble naturelle pour la programmation de jeux vidéo où les coroutines peuvent devenir des entités à part entière et pourraient ainsi communiquer avec d'autres entités par ces mécanismes d'une manière naturelle.

L'envoi de messages se fait par la fonction (`! <corout> <msg>`) qui s'occupe d'acheminer le message désiré à la coroutine spécifiée. La réception de messages, quant à elle, se fait avec deux fonctions distinctes et une forme spéciale :

- (`? [timeout: <sec>]`) : Réception du premier message disponible avec attente bloquante. Un *timeout* peut être spécifié afin de limiter l'attente faite. Si plusieurs messages sont disponibles, le premier arrivé sera alors retourné.
- (`?? <predicat> [timeout: <secs>]`) : Réception du premier message qui retourne vrai selon le prédicat spécifié. Le prédicat doit prendre un seul argument, un message reçu. Comme pour la fonction `?`, l'attente bloquante peut être interrompue après un délai, si ce dernier est spécifié.
- (`recv (<pattern> <body>) ...`) : Cette forme spéciale permet la réception sélective de messages selon des patrons de filtrage par motifs (*pattern matching*). Ces patrons permettent d'exprimer la forme attendue du message et de lier des variables locales à des parties du message reçu afin de les utiliser dans le corps du patron. Lorsque plusieurs messages reçus peuvent correspondre aux motifs donnés, la sélection du message se fait en choisissant le premier message reçu qui correspond à au moins un patron. Si plusieurs patrons peuvent correspondre à ce message, le premier patron (selon l'ordre de définition) sera utilisé. Ainsi l'ordre de spécification des motifs est important et significatif.

Les valeurs de délai d'attente maximale spécifiées pour la réception de messages impliquent qu'il est possible qu'une réception échoue. Dans un tel cas, une exception

de type `mailbox-timeout-exception` est lancée par la fonction de réception utilisée.

Les patrons utilisés dans le filtrage de motifs de la forme spéciale `recv` diffèrent de ceux dans Termite. Dans notre système, lorsqu'un symbole est précédé d'une virgule (*unquote*), celui-ci est *lié* à la valeur trouvée à cet endroit du motif dans le message reçu. Les formats de motifs reconnus sont les symboles, les mots clés, les caractères, les valeurs booléennes, les nombres, les chaînes de caractères, les listes et les vecteurs. Aussi, il est possible d'ajouter des gardes (filtrage conditionnel au prédicat) en incluant une liste de la forme (`where <pred-expr>`) après le motif de filtrage. Le prédicat utilisé est simplement une expression Scheme. Finalement, un patron spécial permettant d'effectuer des attentes bornées dans le temps est disponible sous la forme (`after <timeout> <body> ...`). S'il est présent, ce patron spécial doit être spécifié *en dernier*. La figure 5.7 illustre plusieurs motifs différents qui peuvent être utilisés.

```
(recv (salut                'got-salut)      ; symbol match
      ("bonjour"           'bonjour)        ; string match
      (1011                 '11-or-1011?)    ; number match
      ((a ,b c)             (string b #\ k)) ; list match
      (#(a b ,c)            c)               ; vector match
      (,any-char            ; guarded match
      (where (char? any-char)) 'got-a-char)
      ((tata 1 #(toto ,x) "titi")           ; complex match
      (where (number? x)) (+ x 1))
      (,anything            anything)        ; can match anything
      (after 2              'timeout!))      ; timeout
```

FIGURE 5.7 – Exemple de patrons pouvant être utilisés dans le filtrage par motifs de la forme spéciale `recv`

Une particularité intéressante de ce filtrage par motifs est reliée au fait qu'il est possible d'utiliser des patrons de manière dynamique, *i.e.* qui n'apparaissent pas dans la forme spéciale `recv`, mais plutôt qui ont été spécifiés grâce à une autre forme spéciale : (`with-dynamic-handlers ((<pattern> <handler-body>) ...) <body>`). Tous les appels à `recv` se trouvant dans le corps (`<body>`) de `with-dynamic-handlers` se trouveront augmentés de ces nouveaux patrons. Il est important de mentionner que ces patrons dynamiques seront toutefois considérés en dernier lieu. Il est donc possible de

factoriser des patrons communs et de les appliquer à tous les appels de la forme spéciale de réception de messages par filtrage de motifs.

Le filtrage par patron présenté se conforme à la sémantique définie par les formes équivalentes dans les systèmes Termite [48] et Erlang [47] par le fait que la sélection du message reçu se fait sur le premier message reçu qui correspond à au moins un patron de filtrage. Ce choix semble raisonnable puisqu’il en résulte que la sélection d’un message est complètement déterministe. En effet, la réception de nouveaux messages (qui est indéterministe dû à l’incertitude qu’apporte la fonction `sleep-for`) n’influence pas le choix du message. Si un message déjà reçu correspond à un patron, il sera choisi peu importe l’arrivée d’un nouveau message. Il en découle que le système de messagerie est robuste et fiable.

Aussi, un mécanisme de liste de diffusion a été inclus. Ce mécanisme simple permet d’enregistrer des coroutines dans une liste de diffusion et, par la suite, d’envoyer à toutes les coroutines inscrites des messages de manière simultanée. Quoi que très primitif, ce système permet d’émuler la base d’un système de programmation réactive [49]. Les fonctions de gestion et d’utilisation de listes de diffusion sont :

- (`subscribe <list-id> <corout>`) : Inscription de la coroutine spécifiée à la liste de diffusion identifiée par le symbole choisi. Si la liste n’existe pas, elle sera créée.
- (`unsubscribe <list-id> <corout>`) : Désinscription d’une coroutine à une liste de diffusion.
- (`broadcast <list-id> <msg>`) : Envoi d’un message à une liste de diffusion.

Un identificateur de liste de diffusion est représenté par une donnée Scheme. L’avantage d’utiliser des données comme identificateur est qu’il est alors possible de créer celui-ci durant l’exécution. L’utilisation de variables Scheme pointant vers une structure de données, plutôt que des données Scheme, offrirait de meilleures performances, mais le dynamisme présent dans notre système ne serait pas disponible.

5.1.6 Autres mécanismes de synchronisation

Le mécanisme de messagerie sophistiqué du système de coroutines permet d'effectuer de manière élégante les synchronisations nécessaires entre les coroutines. Toutefois, certains problèmes pourraient être exprimés plus facilement grâce à des moyens traditionnels de synchronisation. Pour ce faire, des sémaphores [50] ont été ajoutées au système de coroutines développé. Celles-ci permettent donc d'exprimer de manière différente les contraintes de synchronisation qui peuvent être nécessaires pour l'implantation d'un jeu. Les fonctions de manipulation et de création de sémaphores sont données ci-dessous :

- `(new-semaphore <init-value>)` : Création d'une nouvelle sémaphore ayant comme valeur initiale `<init-value>`.
- `(new-mutex)` : Création d'une nouvelle sémaphore ayant comme valeur initiale 1.
- `(sem-locked? <sem>)` : Permet de vérifier s'il reste des ressources disponibles dans une sémaphore. S'il en reste, la fonction retournera vrai. Dans le cas contraire, elle retournera faux.
- `(sem-lock! <sem>)` : Prend une ressource de la sémaphore spécifiée. Si aucune ressource n'est disponible, la coroutine se met en état d'attente bloquante jusqu'à ce qu'une ressource soit libérée.
- `(sem-unlock! <sem>)` : Libère une ressource de la sémaphore spécifiée. Aucune vérification n'est faite pour s'assurer que la coroutine courante possède réellement cette ressource. Si des coroutines sont en attente de cette ressource, la première à s'être mise en attente est alors réveillée et enfilée dans la file de coroutines prêtes.

5.2 Implantation

L'implantation de ce système de coroutines est centralisée sur l'utilisation des fonctions `continuation-capture` et `continuation-return` afin de permettre la conser-

vation de l'état courant d'une coroutine. Un net avantage lorsque le système est directement implanté par l'utilisateur est qu'il répond sur mesure aux demandes de ce dernier. Il peut donc s'exprimer dans une syntaxe simple, tout en conservant un contrôle fin sur le comportement du système. Cette section présente les mécanismes internes du système de coroutines développé, en expliquant non seulement les structures de données utilisées, mais aussi les algorithmes.

5.2.1 Implantation des coroutines

La structure de données des coroutines est implantée en utilisant le système d'objets fourni dans le système Gambit-C. Une version orientée objet est aussi disponible. La structure employée est illustrée dans la figure 5.8. Elle contient toutes les informations essentielles au fonctionnement de celle-ci, dont entre autres sa continuation (`kont`), sa boîte de messagerie, une sauvegarde de l'environnement d'un sous-système de coroutines (`state-env`), *etc.*

```
(define-type corout id kont mailbox state-env
                  sleeping? delta-t msg-lists result)

(define corout-unbound-result (gensym 'corout-unbound-result))

(define (new-corout id thunk)
  (let ((kont (lambda (dummy) (terminate-corout (thunk))))
        (mailbox (new-queue))
        (state-env #f)
        (sleeping? #f)
        (delta-t #f)
        (msg-lists (empty-set))
        (result corout-unbound-result))
    (make-corout id kont mailbox state-env
                  sleeping? delta-t msg-lists result)))
```

FIGURE 5.8 – Structure de données représentant une coroutine

La continuation primordiale d'une coroutine est sa terminaison de manière « propre », *i.e.* en utilisant la fonction de terminaison de coroutines. C'est ce qui permet à une coroutine ayant comme *thunk* uniquement `(lambda () 1)` de terminer correctement avec

la valeur de retour 1.

La variable d'état `sleeping?` permet d'indiquer à l'ordonnanceur si la coroutine qui vient de céder le contrôle a été mise en veille ou non, afin de savoir si cette dernière doit retourner dans la file d'attente des coroutines prêtes. La figure 5.3 illustre les états possibles contenus dans `sleeping?`.

Afin de déterminer si la coroutine possède ou non une valeur de retour, un symbole unique connu est utilisé pour initialiser la valeur de retour d'une coroutine. Lorsque l'utilisateur demande à obtenir la valeur de retour d'une coroutine, le contenu du membre `result` est comparé à ce symbole afin de savoir si une valeur de retour s'y trouve ou non, tel qu'illustré dans la figure 5.9.

```
(define (corout-get-result c)
  (if (eq? (corout-result c) corout-unbound-result)
      (raise 'coroutine-not-terminated-exception)
      (corout-result c)))
```

FIGURE 5.9 – Obtention de la valeur de retour d'une coroutine

5.2.2 *Timers*

Les *timers* sont aussi implantés comme des structures *define-type*. Leur rôle est de fournir une abstraction temporelle pour le déroulement du système de coroutines, qui peut être perçu comme une simulation. Cette classe très simple contient des champs afin de tenir compte du temps courant de la simulation, de la période du *timer*, etc.

Les *timers* doivent être rafraîchis régulièrement selon une période fixe. Ainsi, ils doivent être exécutés complètement à l'extérieur du système de coroutines pour y arriver et sont donc implantés en utilisant le système de *threads* de Gambit-C. Il en résulte que ces derniers sont rafraîchis régulièrement de manière concurrente avec le système de coroutines.

5.2.3 Ordonnancement

L'ordonnancement est le coeur du système de coroutines. Cet ordonnancement évolue dans un environnement contenant l'état du système de coroutines actif. Cet environnement est conservé dans une structure de données globale et peut être accédé en appelant la fonction ayant le nom du paramètre. Par exemple, le paramètre d'environnement `current-corout` peut être accédé et modifié par la fonction (`current-corout [<new-value>]`). Le choix de cet *API* a été fait afin de ressembler aux variables dynamiques de Gambit-C. Les paramètres d'états de l'ordonnanceur sont :

- `current-corout` : Paramètre contenant la coroutine actuellement exécutée. Lorsque celle-ci termine son exécution, sa valeur de retour doit être placée dans ce paramètre pour signaler à l'ordonnanceur la terminaison de la coroutine.
- `q` : File d'attente des coroutines prêtes à être exécutées (*ready queue*).
- `timer` : *Timer* utilisé pour la simulation.
- `time-sleep-q` : File prioritaire implantée avec un arbre rouge-noir qui contient les coroutines en attente sur le temps.
- `root-k` : Continuation primordiale, *i.e.* continuation du système de coroutines courant.
- `parent-state` : Sauvegarde de l'état du système de coroutines *parent* au système actuel, pour des systèmes cascades. Cet état est décrit par les paramètres présentés ici.
- `dynamic-handlers` : Liste de patrons dynamiques utilisés avec la forme spéciale `recv`.
- `sleeping-coroutines` : Nombre de coroutines en attente. Cette variable est nécessaire parce que l'accès aux files `q` et `time-sleep-q` n'est pas suffisant, puisque l'information n'est pas centralisée. En effet, il est possible que des coroutines soient

en attente sur une sémaphore ou sur la réception d'un message. Ainsi ce paramètre apporte les informations nécessaires à l'ordonnanceur afin qu'il sache s'il existe toujours au moins une coroutine en attente.

La récursion du système fonctionne grâce à la sauvegarde de l'état du système actuel dans le champ **state-env** d'un objet de coroutine et grâce à la restauration de l'état de la coroutine parent située dans le paramètre d'environnement **parent-state**.

L'algorithme d'ordonnancement, en lui-même, est simple. Ce dernier est présenté intégralement dans la figure 5.10. Dans un premier temps, la coroutine ayant suspendu son exécution est traitée. Si cette dernière a terminé son exécution, sa valeur de retour est sauvegardée, sinon elle est automatiquement remise dans la file d'attente des coroutines prêtes. Par la suite, la file prioritaire de coroutines en attente sur le temps est regardée afin de réveiller toute coroutine ayant dépassé son délai de sommeil. Finalement, la première coroutine disponible dans la file de coroutines prêtes est choisie comme la prochaine et son travail est poursuivi par un appel à **resume-coroutine**. Si toutefois aucune ne se trouvait dans la file **q**, alors une vérification des coroutines en attente sur le temps est faite afin de déterminer s'il reste du travail à faire. S'il y a des coroutines en attente sur le temps, alors l'ordonnanceur se met en veille pour le délai d'attente restant. Cette particularité se distingue nettement des simulations à événements discrets qui auraient plutôt incrémenté leur horloge interne directement de ce délai, comportement qui est indésirable pour un jeu vidéo. Lorsqu'aucune coroutine n'est prête ou ne dort sur le temps, alors la simulation ne peut plus continuer. Dans le cas où il existe d'autres coroutines en veille (par exemple sur l'attente d'un message), alors le système est en position d'interblocage. Sinon, le travail est terminé et donc, l'état du système parent est rétabli et la continuation primordiale est invoquée.

Le changement et le retour de contexte d'une coroutine sont faits par les fonctions **yield** et **resume-coroutine**, illustrées respectivement dans les figures 5.11 et 5.12. La figure 5.11 illustre aussi la version permettant le changement de contexte d'un système

```

(define (corout-scheduler)
  (manage-return-value)
  (wake-up-sleepers)
  (current-corout (dequeue! (q)))
  (cond
    ((corout? (current-corout)) (resume-coroutine))

    ((not (time-sleep-q-empty? (time-sleep-q)))
     (begin
       (let* ((next-wake-time
                (time-sleep-q-el-wake-time
                 (time-sleep-q-peek? (time-sleep-q))))
              (thread-sleep! (/ (- next-wake-time (current-sim-time))
                                (timer-time-multiplier (timer)))))
         (current-corout ___scheduler-is-sleeping___)
         (corout-scheduler)))

    (else
     (let ((finish-scheduling (root-k))
           (ret-val (void)))
       (if (> (sleeping-coroutines) 0)
         (error "Deadlock detected in coroutine system..."))
       (restore-state (parent-state))
       (continuation-return finish-scheduling ret-val))))))

```

FIGURE 5.10 – Algorithme d'ordonnancement

de coroutines récursif. Ces dernières illustrent la base même du système de coroutines, où la réification de continuation permet la sauvegarde de l'état présent du calcul pour une utilisation future. Comme l'indique la figure 5.12, ces calculs sont poursuivis simplement par un retour à leurs continuations sauvegardées dans la structure de données des coroutines. La valeur utilisée pour le retour à ces continuations est complètement arbitraire et sera ignorée par celles-ci.

```

(define (yield)
  (continuation-capture
    (lambda (k)
      (corout-kont-set! (current-corout) k)
      (corout-scheduler))))

(define (super-yield)
  (continuation-capture
    (lambda (k)
      (let ((parent (parent-state)))
        (if parent
          (let ((state (save-state)))
            (restore-state parent)
            (let ((current-c (current-corout)))
              (corout-state-env-set! current-c state)
              (corout-kont-set! current-c k)
              (corout-scheduler)))))))))

```

FIGURE 5.11 – Fonctions de changement de contexte explicite

```

(define (resume-coroutine)
  (let ((kontinuation (corout-kont (current-corout))))
    (if (corout-state-env (current-corout))
      (let ((state (save-state)))
        (restore-state (corout-state-env (current-corout)))
        (parent-state state)))
      (if (procedure? kontinuation)
        (kontinuation 'go)
        (continuation-return kontinuation 'go))))

```

FIGURE 5.12 – Implantation du retour au contexte d'une coroutine

5.2.4 Système de messagerie

La fonctionnalité de messagerie du système de coroutines est très simple. Comme l'illustre la figure 5.8, chaque coroutine possède une boîte de réception de messages. Cette boîte est implantée comme une file d'attente, similairement à la file d'attente des coroutines prêtes utilisée par l'ordonnanceur.

L'envoi de messages consiste à ajouter un nouveau message dans cette file d'attente. Après cet ajout, le système vérifie si la coroutine réceptrice était en attente d'un message et si c'est le cas, alors cette dernière est remise en action dans l'ordonnanceur. La figure 5.13 illustre ce procédé d'acheminement de messages. Puisqu'il est possible de spécifier une valeur d'attente maximale de messages, la coroutine en attente pourrait se retrouver dans la file d'attente sur le temps des coroutines. Afin de distinguer une telle attente bornée d'un appel à `sleep-for`, l'état `interruptible?` du champ `sleeping?` d'une coroutine est utilisé, tel que schématisé dans la figure 5.3.

```
(define (! dest-corout msg)
  (enqueue! (corout-mailbox dest-corout) msg)
  (cond ((sleeping-on-msg? dest-corout)
        (corout-set-sleeping-mode! dest-corout #f)
        (corout-enqueue! (q) dest-corout))
        ((and (sleeping-over-time? dest-corout)
              (interruptible? dest-corout))
         (time-sleep-q-remove! (sleeping-over-time?->node dest-corout))
         (corout-set-sleeping-mode! dest-corout #f)
         (corout-enqueue! (q) dest-corout))))
```

FIGURE 5.13 – Envoi de messages entre coroutines

Pour la réception de messages, la coroutine actuelle vérifie si un message est disponible dans sa boîte de réception et, si c'est le cas, elle retourne le premier message enfilé dans celle-ci. Lorsqu'aucun message n'est disponible, la coroutine attend alors jusqu'à la réception d'un nouveau message avec une attente bornée dans le temps ou pas. Pour une attente bornée, la coroutine est mise en veille pour la durée désirée, en spécifiant qu'elle peut être interrompue par la réception d'un message. Sinon, l'état de la coroutine est alors conservé et celle-ci est considérée comme dormante en attente de

message. Si le délai d'attente est dépassé, alors une exception est lancée à l'utilisateur. La figure 5.14 illustre l'implantation de la fonction `?`, la plus simple pour la réception de messages. Elle permet par contre de donner une bonne idée du procédé employé.

```
(define (? #!key (timeout 'infinity))
  (let ((mailbox (corout-mailbox (current-corout))))
    (if (empty-queue? mailbox)
        (if (not (eq? timeout 'infinity))
            (sleep-for timeout interruptible?: #t)
            (continuation-capture
              (lambda (k)
                (let ((corout (current-corout)))
                  (corout-kont-set! corout k)
                  (corout-set-sleeping-mode! corout (sleeping-on-msg))
                  (corout-scheduler))))))
        (if (empty-queue? mailbox)
            (raise mailbox-timeout-exception)
            (dequeue! mailbox)))))
```

FIGURE 5.14 – Réception de messages avec la fonction `?`

L'implantation de la forme spéciale de réception de messages `recv` est complexe et ne sera pas expliquée en détail. Cette dernière est implantée par une macro Scheme qui transforme les patrons de filtrage donnés dans une deuxième forme spéciale de filtrage par motifs appelée `match`. Le code généré par la macro `recv` tente d'effectuer le filtrage sur chacun des messages présents dans la boîte de réception de la coroutine actuelle. Ainsi, le premier message (dans leur ordre d'arrivée) qui se conforme à un des patrons de filtrage sera retenu. La figure 5.15 illustre ce comportement.

L'expansion macro du corps de la coroutine `c1` est donnée dans la figure 5.16. On constate que le patron donné est vérifié en premier sur chaque message. Par la suite, si aucun message ne correspond à ce patron, une vérification est faite parmi les patrons dynamiques afin de trouver un message pouvant être utilisé. Si aucun message n'est trouvé, alors la coroutine est mise en veille pour une seconde, tel que spécifié par le patron `(after 1 ...)`, et pourrait être interrompue par l'arrivée d'un nouveau message. Il est intéressant de noter que lorsqu'un message est trouvé dans la boîte aux

lettres d'une coroutine, un deuxième appel à la macro `match` est fait afin d'exécuter le corps du patron dans un environnement effectuant les liaisons déclarées dans le motif du patron. Il en résulte ainsi une légère perte de performance, mais l'optimisation de cette double recherche du motif est complexe [51]. Ainsi, puisque l'utilisation de cette forme spéciale a été mise de côté vers la fin de la recherche effectuée pour ce mémoire, cette optimisation a été omise pour des fins de simplicité.

```
(let* ((c1 (new-corout 'c1 (lambda ()
                           (let loop ()
                             (recv
                              (ping (display 'ping-) (loop))
                              (pong (display 'pong-) (loop))
                              (after 1 (display 'finished!)))))))
      (c2 (new-corout 'c2 (lambda () (! c1 'pong))))
      (c3 (new-corout 'c3 (lambda () (! c1 'ping)))))

  (boot (list c2 c3 c1))) => pong-ping-finished!
```

FIGURE 5.15 – Illustration des priorités de réception de messages avec la forme spéciale `recv`

```

(let loop ()
  (let ((#:mailbox44 (corout-mailbox (current-corout)))
        (#:absolute-timeout558 (+ (time->seconds (current-time)) 1)))
    (let #:loop43 ()
      (cond ((queue-find-and-remove!
              (lambda (#:msg45)
                (match #:msg45 (ping #t) (pong #t) (, _ #f)))
              #:mailbox44)
              => (lambda (#:msg46)
                    (match #:msg46
                      (ping (display 'ping-) (loop))
                      (pong (display 'pong-) (loop))
                      (, _ #f)))))

              ((find-value (lambda (pred) (pred)) (dynamic-handlers))
               => (lambda (res) (unbox res) (#:loop43)))

              (else
               (let ((sleep-delta
                     (- #:absolute-timeout558
                        (time->seconds (current-time)))))
                 (if (> sleep-delta 0)
                     (begin
                       (sleep-for sleep-delta interruptible?: #t)
                       (#:loop43))
                     (display 'finished!))))))))

```

FIGURE 5.16 – Expansion macro du corps de la coroutine `c1` de la figure 5.15

5.3 Performances

Afin d’obtenir une idée des performances du système développé, ce dernier est comparé à d’autres systèmes similaires sur les fonctions critiques. La comparaison est effectuée selon le temps (en secondes) pour faire un million de ces opérations. Le choix de celles-ci se résume par les fonctions de changement de contexte (`yield`), d’envoi et de réception de messages simples (`!` et `?`) ainsi que d’envoi et de réception de messages avec filtrage (`!` et `recv`). Les opérations de réception de messages ont aussi été testées en utilisant une attente bornée nulle afin d’observer le coût associé à de telles attentes. Les patrons utilisés pour le filtrage par motifs sans et avec *timeout* utilisent le patron illustré dans la figure 5.17. Finalement, un test plus réaliste d’un serveur de `ping` a été utilisé. Le corps de ce test pour notre système est illustré dans la figure 5.18. Le tableau

Opérations	Coroutines	Gambit-C	Termite	Erlang
yield	2.250	0.385	ND	ND
!	1.155	0.260	0.967	1.282
?	0.728	0.131	5.108	0.270
? avec timeout	1.876	0.142	5.118	0.269
! et ?	6.877	1.328	2.061	1.842
recv	0.684	ND	1.460	0.253
recv avec timeout	2.743	ND	1.938	0.273
ping server	10.796	ND	4.858	2.412

TABLEAU 5.1 – Temps d’exécution (en secondes) de 1 000 000 itérations de tests comparatifs entre le système développé et d’autres systèmes similaires

5.1 présente ces résultats. Le tableau 5.2 présente ces mêmes résultats de manière relative dans le but de mieux illustrer les différences de performances. L’exécution de ces tests a été effectuée sur un processeur de 1.6 GHz possédant 1 Go de mémoire principale.

```
(recv (salut 'non)
      ((allo 12) 'ok)
      <(after 2)>)
```

FIGURE 5.17 – Patron de filtrage utilisé pour les tests de performance. Le message reçu par ce dernier est celui attendu par la deuxième branche

```
(let* ((c2 (new-corout 'c2
  (lambda ()
    (let loop ()
      (recv ((,sender ping) (! sender 'pong)))
      (loop))))))
  (c1 (new-corout
    'c1 (lambda ()
      (do ((i 0 (+ i 1)))
        ((= i benchmark-limit) (kill-all! 'done))
        (! c2 (list (current-corout) 'ping))
        (recv (pong 'ok)))))))
  (time-expr (boot (list c1 c2))))
```

FIGURE 5.18 – Test de performance **ping-server** utilisé pour comparer une utilisation réaliste de la réception de messages avec filtrage de motifs

La comparaison du temps de changement de contexte est uniquement effectuée avec Gambit-C puisque Termite et Erlang ne sont pas des systèmes de coroutines, mais plutôt des systèmes de programmation concurrente. On constate que le changement de

Opérations	Coroutines	Gambit-C	Termite	Erlang
<code>yield</code>	5.8	1.0	ND	ND
<code>!</code>	4.4	1.0	3.7	4.9
<code>?</code>	5.5	1.0	38.9	2.0
<code>? avec timeout</code>	13.2	1.0	36.0	1.8
<code>! et ?</code>	5.1	1.0	1.5	1.3
<code>recv</code>	2.3	ND	5.7	1.0
<code>recv avec timeout</code>	9.8	ND	7.0	1.0
<code>ping server</code>	4.4	ND	2.0	1.0

TABLEAU 5.2 – Comparaison de performances relatives au système le plus rapide (1.0) en fonction des données du tableau 5.1

contexte est environ six fois plus lent que celui effectué par Gambit-C. Puisque le code de `yield` est déjà très épuré, ce test comparatif indique qu’il s’agit plutôt de l’ordonnanceur de notre système qui est plus lent que celui de Gambit-C. Toutefois, cette différence de performance est acceptable pour les utilisations raisonnables de notre système dans un jeu vidéo.

En ce qui concerne l’envoi et la réception de messages simples, notre système se trouve à être environ cinq fois plus lent que le système le plus rapide pour ce genre d’opération, soit Gambit-C. L’utilisation d’attentes bornées par le temps semble toutefois ralentir d’avantage les performances observées, tout comme pour le système Termite. L’écart à Gambit-C se justifie probablement par l’utilisation de structures de données moins efficaces dans notre cas. Malgré tout, les performances observées demeurent acceptables.

La comparaison de filtrage par motif entre notre système, Termite et Erlang indique que notre système et Termite offrent des performances équivalentes, soient environ cinq fois plus lentes que celles offertes par Erlang. Encore une fois, l’utilisation d’attentes bornées a ralenti d’un facteur de deux l’exécution de notre système et celle de Termite. Ces résultats ne sont pas très surprenants puisque Erlang est réputé comme étant très efficace. Toutefois, plusieurs optimisations sur la réception de messages par filtrage pourraient être effectuées dans notre système afin d’améliorer celles-ci.

5.4 Conclusion

Ainsi, un système de coroutines a été implanté afin de fournir aux programmeurs de jeux vidéo une interface à un système permettant d'exprimer des problèmes de changements de contextes (notamment dans le cadre de jeux multi-joueurs), sans avoir à se soucier de synchroniser les coroutines pour éviter les problèmes de sections critiques.

Ce système offre ainsi une interface similaire à celle offerte par le système Ter-mite [46], orientée sur un calcul sériel plutôt que distribué. Il est ainsi possible de concevoir des coroutines comme des entités évoluant dans un même environnement de manière successive, comme c'est souvent le cas dans un jeu vidéo.

Le système a également été généralisé de manière à permettre une utilisation en cascade de systèmes de coroutines. Ainsi, l'utilisation du système ne se limite pas à un usage monolithique, mais permet de séparer des tâches en sous-systèmes de coroutines.

L'utilisation de synchronisation de coroutines par envoi et réception de messages est très naturelle et donc facile à utiliser. Lorsqu'elle est combinée avec des listes de diffusion de messages, on obtient un style de programmation se rapprochant beaucoup des systèmes de programmation réactive [49]. Cette approche est très intéressante pour les jeux vidéo car elle intègre le concept de flot temporel par l'entremise de signaux envoyés entre les entités [52].

Il serait maintenant intéressant d'ajouter un mécanisme de profilage de coroutines au système développé. Un tel mécanisme permettrait d'avoir une meilleure idée du temps moyen que prend une coroutine donnée avant de céder le contrôle dans le but de mieux balancer ou d'optimiser celle-ci.

CHAPITRE 6

ÉVALUATION ET EXPÉRIENCES

Afin de déterminer les forces et les faiblesses du développement de jeux vidéo en Scheme, des jeux doivent être créés, en augmentant graduellement la complexité de ceux-ci, pour permettre de trouver et résoudre de manière itérative les problèmes reliés à leur développement. Un jeu simple comprend les problèmes les plus fondamentaux qui peuvent être rencontrés, soient une détection de collisions, des animations, un concept de niveaux, *etc.* Ainsi, ces problèmes peuvent être adressés dans un premier temps, puis de nouveaux problèmes peuvent être entrepris par la suite en développant un jeu plus complexe. Aussi, cette approche permet de bâtir une infrastructure de développement de jeux vidéo qui réduit la difficulté du développement de jeux plus complexes.

Le jeu choisi pour la première phase de développement est *Space Invaders* (1978). Ce dernier date de la période des jeux d'arcade et, tout en étant très simple, fait face à plusieurs problèmes fondamentaux qu'impliquent les jeux vidéo modernes : des interactions rapides avec usager, des niveaux, de la détection de collisions et même des parties multi-joueurs. Puisque le graphisme de ce jeu est rudimentaire, il constitue un très bon choix pour un premier jeu, car il permet de concentrer le développement sur le moteur de celui-ci. Le développement de ce jeu est traité dans la section 6.1. Son développement a mené à la création du système orienté objets décrit dans le chapitre 4 et du système de coroutines présenté dans le chapitre 5.

Par la suite, le jeu *Lode Runner* (1983) a été développé. Ce dernier date de l'époque des premiers ordinateurs personnels et est plus élaboré. Par contre, c'est la version arcade du jeu (1984) qui a été reprise. Tout en conservant les problèmes fondamentaux traités dans *Space Invaders*, ce jeu étend le concept de niveaux déjà traités, introduit la notion d'intelligence artificielle et présente des entités possédant des états plus complexes. De plus, le développement de *Lode Runner* a pour but de consolider les techniques développées pour la création du premier jeu. Le développement de ce jeu est

traité dans la section 6.2. Les outils et techniques déjà développés ont permis d’aisément résoudre les problèmes rencontrés.

Ainsi, ce chapitre a pour but d’expliquer le cheminement du travail nécessaire à l’écriture de ces deux jeux afin de répondre à la problématique de base traitée par ce document.

6.1 Développement de *Space Invaders*

Ce jeu consiste à « sauver la galaxie » en éliminant une armée d’envahisseurs extra-terrestre grâce à un vaisseau spatial équipé de missiles et bénéficiant de la présence de quatre boucliers. Le joueur ne peut que se déplacer de gauche à droite et tirer des missiles. Les ennemis sont placés en formation et se déplacent aussi latéralement jusqu’à ce qu’ils frappent un côté de l’écran. Ils descendent alors d’une rangée, se rapprochant ainsi du joueur. La vitesse de déplacement des ennemis est inversement proportionnelle à leur nombre. Le niveau se termine lorsque tous les ennemis sont détruits. Alors, le prochain niveau, qui est exactement le même que le précédant, débute. La figure 6.1 illustre une capture d’écran du jeu développé.



FIGURE 6.1 – Capture d’écran du jeu *Space Invaders* développé

L’objectif visé en développant *Space Invaders* est simple : écrire un premier jeu

qui permet d'exposer les problèmes fondamentaux reliés au développement de jeux vidéo et résoudre ces problèmes en tentant de tirer profit de la puissance expressive du langage Scheme. Pour y arriver, une approche de développement itératif (en spirale) a été utilisée.

Une première version a été développée dans le but d'avoir l'infrastructure de base du jeu et d'identifier les problèmes rencontrés au cours du développement. Le développement de cette première version est décrit dans la section 6.1.1.

Par la suite, une deuxième version fut écrite afin de répondre à certaines lacunes que présentait la première version du jeu. Pour ce faire, le développement du système objet présenté dans le chapitre 4 fut entrepris. Cette démarche d'amélioration du jeu est décrite dans la section 6.1.2.

Finalement, une version expérimentale fut développée. Cette version tentait d'intégrer une nouvelle approche face au contrôle de flot du jeu. L'idée était de tenter de fusionner le système objet et le système de coroutines afin que tous les objets du jeu soient eux-mêmes des coroutines, en espérant simplifier l'écriture d'entités du jeu. En expérimentant sur le contrôle de flot du jeu, il est possible de démontrer si l'utilisation de Scheme permet de faciliter de telles expérimentations. Cette dernière version du jeu est décrite dans la section 6.1.3

6.1.1 Version initiale

Dans un premier temps, la version arcade du jeu disponible sur l'émulateur MAME [53] fut étudiée afin d'être reproduite le plus fidèlement possible. MAME est un système émulant plusieurs architectures de machines arcades et permet de jouer sur PC aux versions originales de ces jeux. Plusieurs problèmes potentiels ont alors été identifiés :

- Rendu graphique du jeu
- Flot du contrôle logique du jeu
- Animations

- Structures de données
- Détection et résolution de collisions
- Parties multi-joueurs.

6.1.1.1 Rendu graphique

Afin d'effectuer le rendu graphique du jeu, une librairie de lecture et de chargement d'images a dû être écrite. Plutôt que d'utiliser une librairie C déjà existante pour faire ce travail, le choix de l'écriture de cette librairie visait à limiter les dépendances externes et à avoir autant que possible de code en Scheme. Par contre, les librairies SDL [54] et *OpenGL* [55] ont été utilisées afin de simplifier la gestion d'entrées/sorties, le fenêtrage et le rendu de l'application. L'utilisation des interfaces aux fonctions étrangères (*Foreign Function Interface*) du système Gambit-C ont fait en sorte que la tâche d'accès et d'utilisation de ces librairies fut triviale.

Cette librairie de lecture et de chargement d'images s'occupe de lire des images en format textuel *ppm*, qui est l'un des formats les plus simples pour décrire une image *bitmap*. Le système s'attend à lire des images sous forme de « fontes », *i.e.* contenant plusieurs sous images. Ce choix semble raisonnable car une entité de jeu vidéo est souvent représentée par plusieurs images. Ceci permet donc de garder ces dernières dans une même image. Un fichier descriptif, sous forme de S-expressions, doit être présent afin de permettre une interprétation correcte de la fonte lue. Par exemple, l'image fournie dans la figure 6.2 doit être accompagnée d'un fichier de même nom, portant une extension *.scm* contenant la description `((colors: (white green red)) (chars: (0 1)))` qui définit symboliquement les couleurs de chacune des sous fontes présentes dans l'image ainsi que chacun des caractères présents (qui sont pour cet exemple les états 0 et 1 de l'animation illustrée).

Cette fonte peut alors être lue et chargée en mémoire vidéo grâce à la forme spéciale :



FIGURE 6.2 – Exemple de fonte utilisée dans *Space Invaders*

```
(define-uniform-font <font-name> <width> <height>
  [static] [loop-x] [loop-y])
```

Cette dernière effectue la génération de code C permettant la lecture et le chargement de la fonte spécifiée, à condition que celle-ci soit uniforme, *i.e.* que toutes les sous images soient de même taille. L'option **static** permet d'effectuer la lecture durant l'expansion macro et d'inclure l'image dans le code source généré. Les options **loop-x** et **loop-y** permettent de spécifier le comportement à entreprendre lorsque l'image est agrandie. Par défaut, l'image est lue dynamiquement et est étirée lorsqu'elle est agrandie, mais elle peut être lue dynamiquement et/ou répétée lorsque agrandie dans la direction de l'axe des X ou des Y. Cette librairie est très simple puisqu'elle ne traite que des fontes aux sous images de tailles uniformes. Ce choix fut effectué puisqu'il était bien adapté aux jeux développés et dans le but de ne pas trop complexifier la tâche de la création des jeux.

6.1.1.2 Logique du jeu

Le flot de contrôle de la logique de ce jeu fut expérimental dès le départ. Dans la première version développée, une simulation par événements discrets a été utilisée. L'idée derrière ce choix provenait du fait que dans un jeu vidéo, la progression d'entités peut être considérée comme une série d'événements discrets. Par exemple, le laser avance, le joueur se déplace à droite, le vaisseau ennemi explose, *etc.* Ainsi, un petit système de simulation par événements discrets a été développé et utilisé. Ce dernier fut implanté de manière très simple en utilisant un monceau triant les événements ordon-

nancés de manière croissante en fonction du temps de leur arrivée dans la simulation. Ces évènements étaient des *thunk*, fermetures Scheme sans arguments, effectuant le travail à accomplir par l'évènement. Lorsqu'un évènement se termine, l'ordonnanceur choisit alors le prochain et appelle le corps de celui-ci. Il est également possible pour un évènement d'ordonnancer un autre évènement grâce à la forme spéciale (`in <secs> <thunk>`) qui enregistre cette nouvelle fermeture dans le système de manière à ce qu'elle se fasse appeler dans `<secs>` secondes. Contrairement aux simulations à évènements discrets habituelles, les attentes du système entre les évènements sont effectuées par l'ordonnanceur de manière à respecter les délais demandés par les évènements.

La figure 6.3 illustre des évènements qui sont ordonnancés au tout début d'une partie. Ainsi, le premier évènement se chargera de faire un *dispatch* d'autres évènements débutant la partie. Les deux derniers évènements sont des évènements de gestion d'affichage et d'entrées/sorties du jeu.

```
(schedule-event! sim 0
  (lambda () (new-player! level)
    (in 0 (create-init-invader-move-event level))
    (in 1 (create-invader-laser-event level))
    (in (mothership-random-delay)
      (create-new-mothership-event level))))

(schedule-event! sim 0 (create-main-manager-event level))

(schedule-event! sim 0 (create-redraw-event level))
```

FIGURE 6.3 – Exemple de code de simulation par évènements discrets

Ces évènements doivent donc s'occuper de poursuivre le calcul du jeu leur étant assigné en réordonnant de nouveaux évènements ultérieurement dans le jeu. Par exemple, la figure 6.4 illustre l'évènement principal implantant le comportement désiré pour le vaisseau mère des ennemis.

Cet exemple illustre bien deux désavantages provenant de l'utilisation d'évènements discrets pour implanter le comportement d'entités dans un jeu. Le premier désavantage

```

(define (create-mothership-event level)
  (define mothership-event
    (synchronized-event-thunk level
      (let ((mothership (level-mothership level)))
        (if mothership
          (let ((collision-occured? (move-object! level mothership)))
            (if (or (not collision-occured?)
                    (is-explosion? collision-occured?)
                    (eq? collision-occured? 'message))
              (in mothership-update-interval mothership-event))))))
    mothership-event)

```

FIGURE 6.4 – Évènement représentant le vaisseau ennemi de type *mothership*

est que la plupart des évènements sont récursifs, et donc ils se réordonnent par eux-mêmes un peu plus tard dans le temps. Ceci résulte en une écriture moins intuitive de ce comportement. Aussi, lorsqu'un évènement est déjà prévu, il est possible que l'état du jeu change entre temps rendant cet évènement désuet. Pour le vaisseau mère, il est possible que son prochain déplacement ait été déjà prévu, mais que ce dernier a explosé suite à la collision avec un laser du joueur. De tels évènements sont donc prompts à donner des erreurs dues à des inconsistances entre l'état du jeu attendu et l'état réel lorsque se produit un évènement.

Aussi, il est important de faire bien attention à doser le calcul effectué par un évènement de manière à ne pas monopoliser le temps du processeur utilisé. Ainsi, les évènements sont souvent réordonnés avec un intervalle de temps de zéro, de manière à laisser la chance aux autres évènements de s'exécuter, tout en forçant l'évènement courant de se poursuivre le plus tôt possible.

Toutefois, cette approche semble avoir été très bénéfique pour la création d'animations. En effet, en combinant des évènements discrets avec un style d'écriture en CPS, il est possible de bien modulariser des animations dans le jeu. Par exemple, l'animation du début de jeu peut être écrite sous forme CPS de manière à ce que la suite de cette animation puisse être n'importe quel autre évènement. La figure 6.5 illustre une partie de l'implantation de l'animation de début de partie. Celle-ci reçoit une continuation en paramètre. Elle termine son animation en utilisant une animation CPS s'occupant

de faire clignoter le score du joueur. Puisqu'il s'agit de la fin de l'animation de début de partie, la continuation est passée directement en paramètre à cette animation de clignotement (optimisation d'appel terminal). On a pu ainsi facilement utiliser des animations modulaires dans *Space Invaders*. Cette figure illustre un exemple d'utilisation de cette animation en réécrivant de manière plus complète l'évènement de début de partie de la figure 6.3.

```
(define (start-of-game-animation-event level continuation)
  (lambda ()
    ...
    (in 0 (create-text-flash-animation-event level
      (level-get level score-msg-obj)
      animation-duration new-cont))))

(schedule-event! sim 0
  (start-of-game-animation-event level
    (generate-invaders-event level
      (lambda ()
        (new-player! level)
        (in 0 (create-init-invader-move-event level))
        (in 1 (create-invader-laser-event level))
        (in (mothership-random-delay)
          (create-new-mothership-event level)))))))
```

FIGURE 6.5 – Exemple d'animation combinant les évènements discrets et une programmation CPS

6.1.1.3 Structures de données

Les structures de données utilisées dans *Space Invaders* furent définies par la forme spéciale `define-type` de Gambit-C. Elles ont utilisé l'héritage simple que permet cette forme de manière à tirer profit au maximum du polymorphisme et de la modularisation de code qu'il était possible. La figure 6.6 donne certaines définitions des structures employées.

L'avantage principal de l'utilisation de ces structures est qu'elles sont très performantes (voir le tableau 4.3). Par contre, elles ne permettent pas d'avoir des champs


```
(define-type game-object id type pos state color speed
  extender: define-type-of-game-object)

(define-type-of-game-object invader-ship row col)
(define-type-of-game-object player-ship)
(define-type-of-game-object message-obj text)
...
```

FIGURE 6.6 – Structures de données utilisées dans la première version de *Space Invaders*

communs à toutes les instances d'un type donné. Il en résulte que l'expression du fait que toutes les instances du type **player-ship** doivent être associées à une boîte englobante (*bounding box*) de 13 par 8 pixels est difficile à exprimer. Ces informations pourraient être ajoutées dans chaque instance, mais il en résulterait un gaspillage important d'espace mémoire puisque ces données sont communes à toutes les instances de ce type. Ainsi, pour cette version de *Space Invaders*, un système manuel de type *très rudimentaire* a été utilisé. Une structure de données décrivant un type est créée. Cette dernière contient les données communes aux instances de ce type. Un pointeur vers le bon type est par la suite ajouté dans le champ **game-object-type** de chacune des instances du jeu. L'utilisation de ce petit système de type demande ainsi une écriture de code fastidieuse qui ne devrait pas être nécessaire.

6.1.1.4 Détection et résolution de collisions

La détection de collisions sur ces objets a pu être faite de manière efficace et modulaire grâce aux types qui contiennent de l'information sur les boîtes englobantes des objets. Puisqu'il n'y a jamais beaucoup d'objets présents dans le jeu en même temps, une détection très rudimentaire de collisions fut utilisée. Cette dernière est présentée dans la figure 6.7.

Par contre, un problème majeur relié à l'utilisation de ces structures de données est apparu lorsque la résolution de collisions a dû être implantée. En effet, cette dernière dépend du type des *deux* objets entrant mutuellement en collision. La figure 6.8 illustre bien le problème rencontré. On constate que les types des objets doivent être analysés

```

(define (detect-collision? obj level)
  (or (exists (lambda (collision-obj)
                (if (shield? collision-obj)
                    (obj-shield-collision? obj collision-obj)
                    (obj-obj-collision? obj collision-obj))))
      (level-all-objects level))
      (obj-wall-collision? obj (game-level-walls level))))

```

FIGURE 6.7 – Détection de collision dans *Space Invaders*

manuellement de manière à utiliser la bonne fonction de résolution. Cette dernière se doit aussi d'analyser le type de l'objet reçu de manière à choisir la bonne résolution à adopter. Il en résulte du code très lourd, où l'introduction d'erreurs est facile.

```

(define (resolve-collision! level obj coll-obj)
  (cond
    ((player-ship? obj) (resolve-player-collision! level obj coll-obj))
    ((laser-obj? obj) (resolve-laser-collision! level obj coll-obj))
    ...))

(define (resolve-laser-collision! level laser-obj collision-obj)
  (cond ((invader-ship? collision-obj) ...)
        ((laser-obj? collision-obj) ...)
        ((player-ship? collision-obj) ...)
        ...))

```

FIGURE 6.8 – Résolution de collisions dans la première version de *Space Invaders*

6.1.1.5 Parties multi-joueurs

Les parties multi-joueurs de *Space Invaders* sont très simples. Chaque joueur joue en alternance jusqu'à ce que son vaisseau se fasse détruire. Ainsi, cela implique de conserver deux parties du jeu en *parallèle*, où une seule des deux avance à la fois. Avec un flot de contrôle sous forme de simulation par événements discrets, cela implique d'avoir deux simulations exécutées en parallèle. Puisque le système de simulation ne permet pas d'arrêter une simulation et d'en avoir une deuxième se produisant en même temps, une autre approche a été utilisée. Un système de coroutines dans lequel les simulations sont exécutées à l'intérieur de coroutines différentes a été développé et utilisé. Ainsi, le

changement de contexte peut être manuellement utilisé lorsqu'un joueur meurt afin de réactiver la partie du prochain joueur.

C'est de cette idée qu'est né le système de coroutines présenté dans le chapitre 5. Il fut initialement beaucoup plus simple que celui présenté dans ce chapitre, mais les composantes de base restent les mêmes. La figure 6.9 résume la procédure utilisée lors de la mort d'un joueur. Après avoir effectué la gestion des objets dans le niveau, un mutex est bloqué afin de permettre l'arrêt des animations du jeu, chose requise pour l'animation de mort d'un joueur. Par la suite, l'animation de mort d'un joueur est lancée, avec `continuation` pour continuation. Cette dernière, dans le cas où le joueur n'a pas terminé sa partie, s'occupe d'informer la coroutine de l'autre de certaines informations et d'effectuer le changement de contexte. Il est intéressant de noter que juste avant ce changement, un évènement très prioritaire doit être ordonnancé de manière à ce qu'il survienne immédiatement au retour du contexte de la coroutine. Il s'agit de l'animation ré-introduisant la partie du joueur actuel.

```
(define (explode-player! level player)
  (level-loose-1-life! level)
  (level-add-object! level player-expl-obj)
  (level-remove-object! level player)
  (let ((continuation
        (if (<= (game-level-lives level) 0)
            (begin ...)
            (lambda ()
              (if (2p-game-level? level)
                  (begin
                     (send-update-msg-to-other level #f)
                     (in NOW! (start-of-game-animation-event
                               level (return-to-player-event level)))
                     (yield-corout))
                  (begin ...)))))))
    (in 0 (lambda ()
            (sem-lock! (level-mutex level))
            (in 0 (player-explosion-animation-event
                  level expl-obj animation-duration continuation))))))
```

FIGURE 6.9 – Évènement de mort d'un joueur

6.1.1.6 Conclusion

Ainsi, l'écriture de cette première version de *Space Invaders* a permis d'identifier plusieurs problèmes reliés au développement de jeux vidéo (détection et résolution de collisions, parties multi-joueurs, *etc.*). En plus de permettre d'expérimenter sur le type de flot de contrôle utilisé pour développer ce jeu, Scheme a permis de faire rapidement plusieurs outils non triviaux pour l'implantation du jeu, dont un système de fontes et un petit système de coroutines.

Par contre, certains problèmes ne possèdent pas des solutions satisfaisantes. En effet, la résolution de collisions est très fastidieuse et aurait besoin d'être modularisée et améliorée. Aussi, la technique de flot de contrôle expérimentale utilisée (simulation par évènements discrets), fonctionne bien pour certains aspects tels que la description d'animations, mais ne semble pas naturelle pour d'autres. Traditionnellement, une boucle de jeu (*game loop*), conserve un contrôle étroit sur le contrôle de flot du jeu, mais implique que le comportement des entités sont directement dépendant de cette façon de procéder. Ainsi, le contrôle de flot pourrait aussi être amélioré de manière à exprimer naturellement le comportement des entités du jeu, soit en utilisant une approche traditionnelle ou en expérimentant de nouveau sur une méthode de gestion de contrôle de flot qui serait plus intuitive.

6.1.2 Version orientée objets

Une nouvelle version de *Space Invaders* fut développée afin de pallier à une lacune existante dans la première version écrite : le problème relié aux types de structures de données qui engendrait des cascades fastidieuses de vérifications de types pour la résolution de collisions (voir la figure 6.8).

Initialement, plusieurs systèmes objets existant pour Gambit-C furent mis à l'essai pour pallier à ce problème, mais aucun ne répondait bien aux besoins présents. Certains offraient beaucoup de puissance expressive, mais avec un coût de performance trop élevé tandis que d'autres, plus performants, étaient trop restrictifs dans l'interface offerte.

Pour ces raisons, un système de programmation orientée objet fut développé de manière à bien répondre aux besoins de *Space Invaders*, sans apporter un coût d'utilisation trop élevé. Le développement de ce système fait l'objet du chapitre 4.

Dans cette nouvelle version du jeu, les objets sont ainsi déclarés comme des classes appartenant à une hiérarchie multiple tirant profit pour apposer des propriétés aux objets du jeu. La figure 6.10 illustre un échantillon de la hiérarchie de classes utilisée dans le jeu. On constate que les champs communs à toutes les instances ont été ajoutés comme membres de classe. Une classe abstraite, `sprite-obj`, a aussi été ajoutée afin de donner la propriété de *sprite* aux objets dont le rendu dépend d'images dans une fonte. La classe `invader-ship` hérite des classes `game-object` et `sprite-obj` tout en ajoutant les deux champs communs à tous les types d'*invader*. Cette dernière est elle-même sous classée par les trois types d'ennemis existant (à l'exception du vaisseau mère qui est traité à part). Puisque chaque type d'*invader* possède sa propre classe, on peut alors spécifier les valeurs des champs de classes pour chacune de celles-ci de manière simple et claire. L'utilisation de ces champs pourra être, par la suite, faite de manière complètement uniforme et transparente. La macro `setup-static-fields` a été écrite afin de permettre d'assigner les champs de classes pour les classes enfants de `game-object`.

```
(define-class game-object ()
  (slot: id) (slot: pos) (slot: state) (slot: color) (slot: speed)
  (class-slot: sprite-id) (class-slot: bbox)
  (class-slot: state-num) (class-slot: score-value))

(define-class sprite-obj ())
(define-class invader-ship (game-object sprite-obj) (slot: row) (slot: col))

(define-class easy-invader (invader-ship))
(define-class medium-invader (invader-ship))
(define-class hard-invader (invader-ship))
(setup-static-fields! easy-invader 'easy (make-rect 0 0 12 8) 2 10)
(setup-static-fields! medium-invader 'medium (make-rect 0 0 12 8) 2 20)
(setup-static-fields! hard-invader 'hard (make-rect 0 0 12 8) 2 30)
```

FIGURE 6.10 – Aperçu de la hiérarchie de classe du jeu

L'attrait principal du développement et de l'intégration du système objet pour la création du jeu réside dans l'utilisation des fonctions génériques. Leur utilisation permet d'implanter des opérations spécifiques aux objets de manière complètement transparente. Il en résulte ainsi que l'écriture de ces parties du code sont plus élégantes et solides face aux changements éventuels des structures de données ou de l'architecture du jeu.

Trois fonctions génériques ont été utilisées pour améliorer *Space Invaders* : **detect-collision?**, **resolve-collision!** et **render** qui effectuent respectivement la détection de collisions, la résolution de collisions et le rendu graphique des objets. La figure 6.11 donne un aperçu de l'utilisation faite de la fonction générique **resolve-collision!**. Des instances de cette dernière sont définies pour les paires de collisions possibles dans le jeu et le système s'occupe de faire automatiquement le choix dynamique de la bonne instance en fonction des objets passés en paramètre. Le code résultant est plus modulaire et clair que celui de la version précédente du jeu.

La deuxième version du jeu apporte ainsi plusieurs améliorations face au développement de jeu vidéo en Scheme grâce à l'utilisation de la programmation orientée objet. La modularité et l'abstraction du code de la première version ont été améliorées grâce à l'utilisation d'une hiérarchie de classes annotée de propriétés et comportant des membres de classes et des fonctions génériques pour effectuer de manière transparente les actions de base sur les instances de ces classes. Par contre, un problème n'a toujours pas été adressé dans cette nouvelle itération du jeu. En effet, le flot de contrôle implanté par une simulation à événements discrets n'a pas été modifié afin de mieux être adapté aux comportements des entités du jeu.

```

(define-method (resolve-collision! level (laser laser-obj)
                                         (inv invader-ship))
  (invader-ship? inv)
  (level-increase-score! level inv)
  (destroy-laser! level laser)
  (explode-invader! level inv))

(define-method (resolve-collision! level (laser1 laser-obj)
                                         (laser2 laser-obj))
  (let ((inv-laser (if (player-laser? laser1) laser2 laser1)))
    (explode-laser! level inv-laser)
    (destroy-laser! level inv-laser)))

...

(define (move-object! level obj)
  (move-object-raw! obj)
  (let ((collision-obj (detect-collision? obj level)))
    (if collision-obj
        (begin (resolve-collision! level obj collision-obj)
                collision-obj)
        #f)))

```

FIGURE 6.11 – Exemple de déclarations d’instances de la fonction générique `resolve-collision!` dans *Space Invaders*

6.1.3 Version avec flot de contrôle sous forme de coroutines

Après deux versions du jeu, *Space Invaders* possède toujours certains problèmes dont celui du flot de contrôle utilisé qui n’est très bien adapté aux objets du jeu. Afin d’adresser ce problème, le système de coroutines utilisé pour l’implantation de parties multi-joueurs a été étendu de manière à ce que le flot de contrôle des entités du jeu soit exprimé plutôt sous forme de *coroutines*. Ce système est présenté dans le chapitre 5. Cette approche est aussi expérimentale, mais l’utilisation d’une simulation à événements discrets a laissée croire que la description du comportement des entités sous forme de *threads* serait plus naturelle. En effet, en utilisant des événements discrets, le corps des événements devait faire en sorte que l’événement se poursuive en se réordonnant un peu plus tard, donnant ainsi la chance aux autres événements de s’exécuter. Ce phénomène correspond très bien à la fonction `yield` du système de coroutine qui interrompt de manière temporaire l’exécution d’une coroutine afin de laisser la chance aux autres

coroutines de continuer leur travail.

Ainsi, dans cette nouvelle version du jeu, *tous* les objets sont maintenant des coroutines exécutant leurs comportements de manière indépendante. Pour ce, une version orientée objet du système de coroutine fut produite. La figure 6.12 illustre l'intégration du système de coroutines aux objets du jeu. L'utilisation de constructeurs permet l'appel du constructeur de base des coroutines afin de spécifier le comportement que devra entreprendre l'objet. Ce corps est déterminé de manière dynamique avec la fonction générique `behaviour` et sera exécuté dans un environnement contenant des gestionnaires de messages dynamiques, communs à tous les objets, permettant de pauser le jeu ou de détruire ceux-ci.

```
(define-class game-object (corout)
  (slot: pos)
  (slot: state)
  (slot: color)
  (slot: speed)
  (class-slot: sprite-id)
  (class-slot: bbox)
  (class-slot: state-num)
  (class-slot: score-value)
  (constructor: (lambda (obj id pos state color speed level)
    (init! cast: '(corout * *) obj id
      (lambda ()(with-dynamic-handlers
        ((pause (pause obj level))
          (die (die obj level)))
        ((behaviour obj level))))))
    (set-fields! obj game-object
      ((pos pos) (state state)
        (color color) (speed speed))))))
```

FIGURE 6.12 – Classe de base des objets en tant que coroutines

Malgré l'idéologie prometteuse d'avoir des objets agissant de manière indépendante, l'implantation du comportement des entités du jeu est devenu très rapidement *infernale*. En effet, puisque toutes les entités possèdent leurs fils d'exécution et que ceux-ci sont tous aussi prioritaires, la *synchronisation* entre ces dernières est primordiale et implique des protocoles de communication complexes entre les objets. Ces protocoles doivent régir

tout le comportent du jeu et, en bout de ligne, le flot de contrôle devient soudainement très difficile à exprimer. Ainsi, des méta-objets ont dû être introduits afin d'effectuer la synchronisation entre les objets de base, comme les ennemis. La figure 6.13 donne le comportement du méta-objet régissant l'activité d'une rangée d'*invaders*. Ce dernier utilise des listes de diffusion pour envoyer des messages aux *invaders* de la rangée qu'il coordonne. Il possède deux états : l'état initial qui attend la réception d'un message lui indiquant que c'est le tour de sa rangée de se déplacer et un deuxième état consistant à une barrière de synchronisation effectuant l'attente de la réception du message `moved` de la part de tous les *invaders* de la rangée. Puis, il vérifie si une collision s'est produite avec un mur durant se déplacement et effectue la gestion de cette collision en prévenant les autres contrôleurs de la situation.

De plus, l'utilisation intensive (voir abusive) du système de coroutine couplé avec le système d'objets a résulté en une chute dramatique des performances du jeu. Cette chute de performance a permis d'améliorer le système de coroutines en implantant un système de profilage minimaliste de coroutines afin de visualiser quelles coroutines monopolisaient le contrôle de l'ordonnanceur pour une durée trop longue. Il fut ainsi possible d'optimiser certaines coroutines afin d'améliorer les performances globales du système. Toutefois, même après avoir effectué ces optimisations, les performances du jeu demeurèrent très mauvaises et donc, la troisième itération de *Space Invaders* fut abandonnée en cours de route. La majeure partie de la logique du jeu fut implantée sous forme de coroutines. Cette version a permis d'illustrer que l'approche plus traditionnelle pour le flot de contrôle serait plus appropriée que les approches utilisées pour implanter *Space Invaders*. Par contre, cette expérimentation sur la technique de flot de contrôle a permis de démontrer la facilité avec laquelle il a été possible de modifier le coeur du jeu pour essayer de nouvelles options quant à l'architecture du programme et de son flot de contrôle.

```

(define (invader-controller)
  (define (init-state)
    (recv
      (init
        (subscribe instant-components (self))
        (broadcast `(invader-row ,(Inv-Controller-row (self)))
                    'move)
        (wait-state))))))

(define-wait-state wait-state moved (inv-nb)
  (recv
    (go-down-warning
      (let ((row-nb (Inv-Controller-row (self))))
        (broadcast `(invader-row ,row-nb) 'wall-collision)
        (for i 0 (< i invader-row-number)
          (if (not (= i row-nb))
              (if (not (zero? (msg-list-size `(invader-row ,i))))
                  (let ((move-back? (< i row-nb)))
                    (broadcast `(invader-row ,i)
                                `(wall-collision ,move-back? #f))))
              (broadcast `(invader-row ,i)
                          `(wall-collision #t ,(self)))))))
    (wait-state))
  (after 0
    (wait-for-next-instant)
    (broadcast `(row-controller
                  ,(next-row (Inv-Controller-row (self))))
              'init)
    (unsubscribe instant-components (self))
    (init-state))))

(init-state))

```

FIGURE 6.13 – Comportement du méta-objet `invader-controller` qui régit le comportement d’une rangée d’*invaders*

6.1.4 Conclusion

Ainsi, le développement de *Space Invaders* en Scheme fut très riche en expériences. Il a permis de trouver plusieurs problèmes reliés au développement de jeux vidéo. Ces derniers ont été résolus en utilisant le mieux possible les avantages qu’offre un langage de haut niveau tel que Scheme fournissant des fonctions d’ordres supérieures, permettant un style de programmation CPS, *etc.*

La création d’un système de programmation orienté objet a grandement permis

d'améliorer la qualité du code écrit lors du développement de la première version du jeu grâce à l'utilisation des fonctions génériques permettant de faire un choix dynamique de procédures en fonction des objets passés en paramètre. Notamment, la résolution de collisions, qui dépend du type des deux objets en question, a pu être écrite de manière très élégante et modulaire grâce aux fonctions génériques.

L'expérimentation effectuée sur le flot de contrôle du jeu a permis de démontrer que l'architecture du moteur peut être modifiée afin de mettre à l'essai de nouvelles techniques de gestion du flot de contrôle. L'utilisation du système de coroutines permis aussi d'implanter efficacement une version multi-joueurs du jeu, mais fut toutefois inappropriée pour gérer efficacement le flot de contrôle du jeu.

Puisque *Space Invaders* est un jeu très simple, la gestion mémoire automatique n'a jamais présenté d'obstacles face au développement puisque trop peu d'objets étaient utilisés et puisque les pauses dues à la récupération mémoire n'utilisaient qu'environ 2 millisecondes, soit approximativement 1 % du temps processeur requis pour le rendu d'une image.

Il serait maintenant intéressant d'implanter un jeu plus complexe afin de valider si les techniques utilisées pour *Space Invaders* sont toujours applicables et de voir si elles s'étendent bien aux nouveaux défis de programmation résultant de l'implantation d'un tel jeu.

6.2 Développement de *Lode Runner*

Après avoir développé *Space Invaders*, un deuxième jeu a dû être développé afin de consolider les techniques développées. Le jeu *Lode Runner* a été choisi dans le but de remplir cette tâche pour plusieurs raisons. Principalement, ce jeu contient de nouveaux éléments qui n'étaient pas présents dans *Space Invaders*, tel que le concept d'intelligence artificielle, et étendait d'autres concepts qui y étaient très primitifs comme le concept de niveaux et d'états d'entités.

Ce jeu consiste à tenter de capturer tout l'or dispersé dans le niveau en évitant

d'entrer en contact avec les robots ennemis. Pour y arriver, le joueur doit grimper des échelles, passer sur des cordes raides et utiliser son pistolet laser pour percer des trous au sol près du joueur. Ces derniers se referment après un délai prescrit et donc le joueur doit se dépêcher afin d'éviter de rester pris au piège à l'intérieur d'un trou. Lorsque tout l'or est en possession du joueur, une échelle de sortie apparaît et mène le joueur au niveau suivant. Le joueur dispose de trois vies pour progresser dans les niveaux du jeu et obtenir le meilleur pointage possible. La figure 6.14 illustre une capture d'écran du jeu développé.

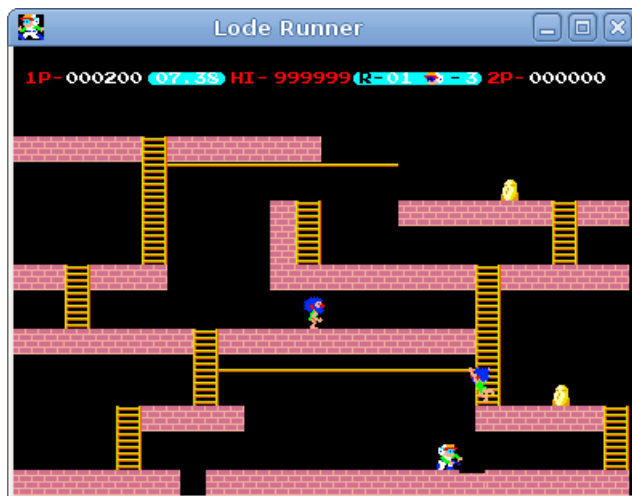


FIGURE 6.14 – Capture d'écran du jeu *Lode Runner* développé

6.2.1 Logique du jeu

Afin de développer ce jeu, les meilleures techniques utilisées pour le développement de *Space Invaders* ont été réutilisées et raffinées pour les besoins plus grands de *Lode Runner*. Par contre, le flot de contrôle de la logique du jeu est régi de manière beaucoup plus conservatrice en utilisant une boucle principale (*Game Loop*). Celle-ci effectue le travail à faire pour chaque image du jeu rendue. Pour ce faire, une fonction générique `advance-frame!` est appelée avant le rendu de chaque image de manière à effectuer le travail que doivent accomplir les objets du niveau en cours. La figure 6.15 illustre une partie de la boucle principale du jeu et l'instance de cette fonction générique pour un

niveau de jeu. Cette figure illustre la simplicité résultante du flot de contrôle du jeu grâce à l'utilisation de fonction génériques et de fonctions d'ordres supérieures.

```
(define-method (advance-frame! (level level) keys-down keys-up)
  (cond ((level-get 'player level) =>
        (lambda (player)
          (player-velocity-set! player point-zero))))
  (cond
    ((level-paused? level) 'do-nothing)
    ...
    (else
     (update! level level current-time (lambda (t) (+ t (fl/ 1. (FPS)))))
     (if (> (level-current-time level) (level-time-limit level))
         (level-game-over! level)
         (begin
          (for-each (flip process-key level) keys-down)
          (for-each (flip animate level) (level-objects level)))))))

;; Game loop
(let loop ((level (current-level)))
  (if exit-requested? (quit))
  (poll-SDL-events)
  (advance-frame! level (key-down-table-keys) (key-up-table-keys))
  (render-scene screen level)
  (loop (current-level)))
```

FIGURE 6.15 – Boucle principale du jeu *Lode Runner*

Par contre, l'utilisation d'une telle boucle fait en sorte que la vitesse du jeu dépend directement du taux de rafraîchissement de celui-ci.

6.2.2 Gestion de l'état des objets

Les entités du jeu *Lode Runner* sont sensiblement plus complexes que celles présentes dans le jeu précédent. En effet, ces nouvelles entités sont animées et réagissent différemment en fonction du contexte de celles-ci. Encore une fois, ce sujet est traité de manière traditionnelle en utilisant des états d'objets et des machines à états afin de régir leurs comportements et leurs animations. La figure 6.16 donne le contenu de la classe `human-like` qui sert de base aux classes `player` et `robot` qui implantent respectivement les instances

du joueur et des robots ennemis. Elle démontre bien la complexité de l'état de ces entités.

```
(define-class human-like (game-object moving statefull)
  (slot: can-climb-up?)
  (slot: can-go-down?)
  (slot: can-use-rope?)
  (slot: can-walk?)
  (slot: dropped-rope?)
  (slot: stuck-in-hole?)
  (slot: facing-direction)
  (slot: walk-cycle-state)
  (slot: shooting?)
  (slot: escaping?)
  (constructor: (lambda (self x0 y0 initial-velocity id) ...)))
```

FIGURE 6.16 – Définition de la classe `human-like` dans le jeu *Lode Runner*

La fonction générique `change-state!` utilise l'état des instances de ces classes pour déterminer l'animation qui doit être utilisée afin d'obtenir un rendu cohérent de l'entité. La figure 6.17 contient la gestion principale de l'état des entités appartenant à la hiérarchie de la classe `human-like`. Ici, avant le rendu de chaque image, l'état de chaque objet est vérifié et mis-à-jour. Puisque les états possibles d'une instance de la classe `human-like` ne sont pas orthogonaux, *i.e.* que plusieurs états correspondant à des animations différentes peuvent être présents en même temps, l'ordre de la modification de l'état est significatif. Si ces derniers avaient été orthogonaux, la gestion de ces états aurait pu être effectuée en utilisant la discrimination de méthodes sur les valeurs de ces états. Le code résultant aurait donc été très limpide.

Les fonctions d'animation des entités (`walk-cycle`, `rope-cycle!`, ...) produisent des animations se répétant de manière cyclique. Puisque ce genre d'animation est fréquent dans le jeu, il fut abstrait dans une forme spéciale afin de simplifier la définition de celles-ci. La figure 6.18 présente cette dernière qui construit une définition de fonction qui met-à-jour les champs `cycle-member` et `state-member` afin de produire correctement le rendu de l'animation. Le champs `cycle-member` de l'instance reçue contiendra un nombre correspondant à l'état actuel de l'animation. Cet état est conservé dans

```

(define-method (change-state! (hl human-like) level)
  (let* ((v (moving-velocity hl)))
    (cond
      ((eq? (human-like-escaping? hl) 'escaping) (ascend-cycle! hl))
      ((human-like-shooting? hl) (shoot-cycle hl))
      ((human-like-stuck-in-hole? hl) (dying-cycle! hl))
      ((and (not (zero? (point-x v)))
            (human-like-can-walk? hl))
       (walk-cycle! hl))
      ((human-like-can-use-rope? hl)
       (if (or (not (zero? (point-x v)))
               (not (memq (human-like-state hl) rope-states)))
           (rope-cycle! hl)
           'keep-same-state-^_^))
      ((not (human-like-can-go-forward? hl)) (fall-cycle! hl))
      ((not (zero? (point-y v))) (ascend-cycle! hl))
      (else (reset-walk-cycle! hl))))))

```

FIGURE 6.17 – Gestion de l'état d'une entité appartenant à la hiérarchie de type *human-like* dans *Lode Runner*

l'attribut `state-member` et sera utilisé par l'algorithme de rendu graphique du jeu. Le paramètre `cycle-delta` représente le nombre d'images qui seront rendues avant le changement de l'état actuel de l'entité.

```

(define-macro (define-cyclic-animation name
  #!key base-class cycle-delta cycle-member
        state-member states other-actions-fun)
  (let ((obj (gensym 'obj))
        (cycling-delta (gensym 'cycling-delta))
        ...)
    `(define (,name ,obj)
      (let* ((,cycle-length (* ,cycle-delta ,(length states))))
        (update! ,obj ,base-class ,cycle-member
          (lambda (s) (modulo (+ s 1) ,cycle-length)))
        (let ((,next-state
              (case (quotient (,cycle-member-getter ,obj) ,cycle-delta)
                ,@(map-with-index (lambda (i x) `((,i) ',x)) states)
                (else (error ...))))))
          (,state-member-setter ,obj ,next-state)
          (,other-actions-fun ,obj))))))

```

FIGURE 6.18 – Forme spéciale utilisée afin de simplifier la définition d'animations cycliques d'entités du jeu *Lode Runner*

Ainsi les animations du jeu peuvent être définies en utilisant cette forme spéciale. La figure 6.19 illustre la définition de l’animation de marche utilisée par les sous-classes de `human-like`. On peut ainsi tirer profit du polymorphisme du système objet en plus du petit langage de définition d’animations afin de définir de manière claire et concise une animation cyclique.

```
(define-cyclic-animation walk-cycle!
  base-class: human-like
  cycle-delta: 5
  cycle-member: walk-cycle-state state-member: state
  states: (standing-up standing-left standing-up standing-right)
  other-actions-fun:
    (lambda (obj) (human-like-facing-direction-set!
                   obj (human-like-get-direction obj))))
```

FIGURE 6.19 – Définition d’animation cyclique dans le jeu *Lode Runner*

En plus de considérer les objets du jeu comme des machines à états, les niveaux eux-mêmes peuvent aussi être considérés comme tels. En effet, on peut subdiviser le flot d’un niveau en six états différents : `pre-game`, `start`, `in-game`, `paused`, `game-over` et `level-cleared`. La figure 6.20 illustre les transitions possibles entre ces derniers. Chaque état est justifié par un comportement différent du niveau, soit par une gestion différente des entrées du joueur comme c’est le cas pour l’état `paused` ou encore par un rendu graphique différent du niveau.

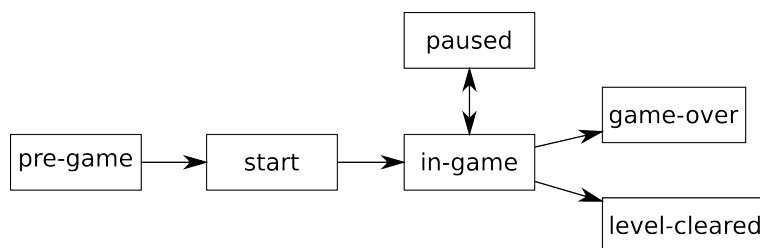


FIGURE 6.20 – Diagramme d’états d’un niveau de *Lode Runner*

Afin d’implémenter ce système de machines à états, un petit langage spécifique au domaine a été développé. Ce dernier tire avantage de la discrimination de méthodes

sur des valeurs du système de programmation orientée objet utilisé afin de valider les transitions effectuées. Sans trop entrer dans les détails, il ne suffit que de sous-classer la classe `state-machine` par les objets de niveaux et ensuite d'effectuer la déclaration de la machine à états faite selon l'interface présentée dans la figure 6.21. Les états sont normalement décrits par des symboles. Les actions de transition peuvent être n'importe quelle expression Scheme. Aussi, lors de la description des transitions, il est possible d'utiliser le symbole `*` afin de représenter tous les états possibles ou sinon d'utiliser une sous-liste contenant les états de provenances ou d'arrivées pour cette transition. La figure 6.22 illustre la définition de la machine à états des niveaux du jeu.

```
(define-state-machine
  <class-name> <init-state> <init-action> <state-list>
  ((<from-state> <to-state> <transition-action>) ...))
```

FIGURE 6.21 – Interface du système de machines à états développé pour les niveaux de *Lode Runner*

```
(define-class level (menu state-machine) ...)

(define-state-machine level
  'pre-game pre-game-transition
  (pre-game start in-game game-over level-cleared paused)
  ((pre-game start      level-start-transition)
   (start    in-game    level-in-game-transition)
   (in-game  game-over  level-game-over-transition)
   (in-game  level-cleared level-cleared-transition)
   (in-game  paused     level-paused-transition)
   (paused  in-game     level-paused-transition)
   (* * (lambda (self) (println "unknown transtion into " to))))
  create-new-class?: #f)
```

FIGURE 6.22 – Implantation de la machine à états pour les niveaux de *Lode Runner*

Les actions utilisées lors des transitions servent principalement à initialiser le nouvel état du niveau. Les fonctions de rendu et de gestion des entrées du joueur s'occupent elles-mêmes de faire la distinction entre ces états. Par exemple, la fonction `level-start-transition` qui sert à faire le compte à rebours du début de la partie s'occupe d'ajouter un nouvel objet représentant une chaîne de caractères qui contient

le compte à rebours. Ce dernier utilise la récursivité afin de se recréer avec le rebours mis-à-jour à toutes les 60 images rendues, tel qu'illustré dans la figure 6.23.

```
(define (level-start-transition lvl)
  (define (make-counter-label counter)
    (if (>= counter 0)
        (let ((lab
              (new label
                'start-counter
                (if (zero? counter) "Start!" (number->string counter))
                24 18 'green
                (list (lifetime 60
                      (lambda (l lvl)
                        (level-delete! l lvl)
                        (make-counter-label (- counter 1)))))))
          (gui-centered?-set! lab #t)
          (gui-scale-ratio-set! lab 4.)
          (level-add! lab lvl))
        (transition lvl 'in-game)))
  (make-counter-label 3))
```

FIGURE 6.23 – Action de transition vers l'état **start** d'un niveau du jeu *Lode Runner*

Ainsi, les macros et les fonctions d'ordres supérieures de Scheme ont permis d'exprimer de manière élégante et modulaire les concepts reliés à la déclaration de machine à états pour les entités ainsi que les niveaux du jeu *Lode Runner*.

6.2.3 Détection et résolution de collisions

Une nette amélioration apportée au jeu, en comparaison de celui développé précédemment, est l'utilisation d'une grille pour accélérer la détection de collisions. Le principe est simple : une grille bidimensionnelle contenant un nombre de subdivisions variable est créée. Lorsqu'un objet est ajouté au jeu, il est ajouté dans chacune des cases de la grille auxquelles il entre en contact et, réciproquement, cet objet possède des pointeurs vers chacune de ces cases de la grille. Si la grille est assez fine, alors la détection de collisions devient triviale : il ne suffit que de regarder les cases de la grille qui touchent l'objet en question afin de voir si d'autres objets s'y trouvent et si c'est le cas, alors une collision est détectée. Bien sûr, la taille de la grille influence le travail nécessaire

lors de chaque ajout et déplacement d'objets. Toutefois, puisque dans *Lode Runner* la plupart des objets sont statiques, on obtient une bonne accélération de la détection de collisions. La figure 6.24 illustre l'algorithme de détection de collisions utilisé.

```
(define (detect-collisions obj level)
  (filter (lambda (x) (not (eq? x obj)))
    (fold-l (curry2* set-union eq?)
      '()
      (map (curry2 grid-get (level-grid level))
        (game-object-grid-cells obj))))))
```

FIGURE 6.24 – Détection de collisions dans *Lode Runner*

La résolution de collisions est effectuée de manière très similaire à ce qui a été utilisé pour *Space Invaders*. Une fonction générique **resolve-collision** s'occupe de choisir dynamiquement la bonne instance de cette fonction générique en discriminant sur les objets concernés par la collision.

6.2.4 Rendu graphique

Il est intéressant de noter que le rendu graphique est légèrement plus complexe que pour le jeu *Space Invaders*, car il est possible d'avoir des objets qui se superposent. Afin de résoudre les problèmes de visibilité qui en résultent, l'algorithme du peintre a été utilisé. Ce dernier est implanté en utilisant des couches sur lesquelles doivent apparaître les objets afin de donner des priorités de rendu. Comme l'indique la figure 6.25, le joueur se trouve sur la couche **human-like-layer** qui est prioritaire face aux couches des objets du jeu, résultant en l'affichage du joueur par-dessus ces derniers. Tous les objets sont triés lorsqu'ils sont ajoutés au niveau afin d'éviter de faire ce tri lors du rendu de chaque image.

```
(enum background-layer stage-layer foreground-layer
  human-like-layer top-layer)

(define-method (get-layer (h human-like))    human-like-layer)
(define-method (get-layer (h hole))           foreground-layer)
(define-method (get-layer (s stage))          stage-layer)
(define-method (get-layer (obj game-object)) foreground-layer)
```

FIGURE 6.25 – Implantation des couches de rendu dans *Lode Runner*

6.2.5 Intelligence artificielle

L'intelligence artificielle du jeu fut implantée de manière directe en utilisant des fermetures afin d'abstraire les différents algorithmes potentiels (proposant différentes difficultés aux joueurs). Seuls deux algorithmes furent toutefois implantés pour le jeu. Le premier, complètement trivial, fait en sorte que les robots demeurent parfaitement immobiles. Le deuxième utilise un environnement totalement observable et reproduit un comportement *similaire* à celui des ennemis de la version originale. Ces algorithmes sont présentés dans les figures 6.26 et 6.27. Les fermetures contenant ces algorithmes sont utilisées de manière dynamique en fonction du niveau de difficulté choisi par le joueur. Il est ainsi possible de modifier durant la partie la difficulté. Aussi, l'utilisation de fermetures donne la possibilité d'en créer de nouvelles contenant de l'information plus spécialisée pour l'état actuel du jeu. L'implantation du choix dynamique de l'algorithme d'intelligence artificielle est présenté dans la figure 6.28.

```
(define (immobile-ai robot level)
  (robot-velocity-set! robot point-zero))
```

FIGURE 6.26 – Algorithme trivial d'intelligence artificielle des robots

```

(define (seeker-ai robot level)
  (let ((player (level-get 'player level)))
    (if player
      (let* ((dir (point-sub player robot))
              (y-axis? (let ((x (point-x dir)) (y (point-y dir)))
                          (and (not (zero? y))
                               (can-go-up/down? robot x y))))
              (velo (if y-axis?
                        (new point 0 robot-movement-speed)
                        (new point robot-movement-speed 0)))
              (factor (if (< (if y-axis? (point-y dir) (point-x dir)) 0)
                          -1
                          1)))
        (robot-velocity-set! robot (point-scalar-mult velo factor))))))

```

FIGURE 6.27 – Algorithme d'intelligence artificielle reproduisant un comportement *similaire* au jeu original

```

(define (get-ai-fun difficulty)
  (case difficulty
    ((easy) immobile-ai)
    ((hard) seeker-ai)
    (else immobile-ai)))

(define (run-ai robot level)
  ((get-ai-fun (level-difficulty level)) robot level))

```

FIGURE 6.28 – Utilisation dynamique des algorithmes d'intelligence artificielle

6.2.6 Conclusion

L'implantation du jeu *Lode Runner* a permis de consolider certaines des techniques d'implantation de jeux développées pour *Space Invaders*. En effet, l'utilisation de fonctions génériques a permis d'exprimer simplement et de manière très *modulaire* le comportement des objets dans le jeu.

Le flot de contrôle du jeu a été implanté de manière beaucoup plus traditionnelle que pour *Space Invaders*. Il en résulte que le jeu dépend beaucoup des notions d'états des objets. Puisque le langage Scheme offre le meilleur des deux mondes, *i.e.* qu'il offre un langage fonctionnel permettant des mutations, ce type de contrôle de flot a permis de tirer profit des avantages qu'offrent la programmation fonctionnelle (fonctions d'ordres supérieures, fermetures, *etc.*) tout en utilisant un modèle éprouvé de flot de contrôle et de modularisation.

La gestion des états plus complexes des objets du jeu a pu être écrite de manière très élégante grâce à la conception d'un langage spécifique au domaine minimaliste. Il fut ainsi possible de décrire de manière déclarative les animations tout en cachant la mécanique commune aux animations des entités du jeu.

L'ajout d'une grille pour accélérer la détection de collisions semble peut-être peu significative puisque *Lode Runner* ne contient pas une très grande quantité d'objets, mais ce système a été ajouté dans le but de facilement et rapidement implanter des jeux qui seront plus complexes et qui nécessiteraient de telles optimisations.

Tout comme ce fut le cas pour *Space Invaders*, la gestion mémoire automatique n'a pas posé de problèmes pour le développement de ce jeu. En effet, malgré un plus grand nombre d'objets présents, ceux-ci demeurent en trop faible quantité pour éprouver la gestion mémoire sur des ordinateurs récents.

6.3 Performances

La performance des jeux développés a été évaluée sur plusieurs aspects. Ces derniers mesurent l'efficacité des jeux sur l'utilisation de ressources (processeur et mémoire),

Points de comparaison	<i>Space Invaders</i>	<i>Lode Runner</i>
Utilisation du processeur	40 %	30 %
Nombre d'images par seconde	29.8	29.0
Durée minimum de pause du GC	0.0020 s	0.0033 s
Durée moyenne de pause du GC	0.0025 s	0.0045 s
Durée maximum de pause du GC	0.0037 s	0.0096 s
Quantité de mémoire utilisée	1 Mo	1.5 Mo

TABLEAU 6.1 – Analyse de performances des jeux développés

ainsi que sur l'impact du système de gestion automatique de mémoire. Le tableau 6.1 illustre les résultats obtenus face à ces points de comparaison. Les résultats présentés dans ce tableau ont été obtenus en jouant durant environ une minute avec les versions compilées des jeux. Il est important de noter que les jeux possèdent un taux de rafraîchissement de 30 images par seconde. Ce taux a été fixé dans le but d'avoir un comportement similaire aux versions originales de jeux implantés.

On constate que le jeu *Space Invaders* utilise d'une manière plus intensive le processeur principal que *Lode Runner* qui pourtant est un jeu plus complexe. Cette perte de performance est attribuée à l'utilisation intensive des systèmes objet et de coroutine développés. En effet, *Lode Runner* a fait une utilisation plus légère de ces systèmes de manière à tirer profit de leurs abstractions aux endroits les plus appropriés. Il en résulte ainsi en une amélioration des performances d'environ 10 %. Une comparaison de ces performances avec les versions originales des jeux est présentée dans le tableau 7.2 du chapitre 7.4.

En ce qui concerne l'utilisation de la mémoire, les deux jeux présentent des profils acceptables. En effet, l'utilisation moyenne de la mémoire varie entre un et deux méga-octets, ce qui correspond à une très faible utilisation de la mémoire sur des plate-formes modernes.

Finalement, on peut observer que les pauses dues à la récupération de mémoire automatique furent de très courte durées. Ainsi, celles-ci n'ont eu aucun impact sur la fluidité du jeu tel que requis. Toutefois, la faible utilisation mémoire vient biaiser cette observation. On ne peut donc pas conclure dans l'absolu sur l'impacte de l'utilisation

de la gestion de mémoire automatique dans un jeu développé en Scheme avec nos implantations de *Space Invaders* et *Lode Runner*.

6.4 Conclusion

Ainsi, il a été possible de développer deux jeux vidéo de complexités croissantes en utilisant le langage Scheme. Le développement du premier jeu, *Space Invaders*, a permis d'identifier et de répondre aux besoins fondamentaux de jeux vidéo : rendu graphique, détection de collisions, *etc.* Pour ce faire, un système de programmation orientée objet et un système de coroutines ont été ajoutés au langage Scheme. Aussi, le premier jeu a permis de vérifier la difficulté inhérente au changement de l'architecture du flot de contrôle du moteur en passant d'un flot contrôlé par une simulation par événements discrets à une simulation par agents. La conversion s'est faite sans trop d'efforts, grâce aux abstractions utilisées.

Par la suite, un deuxième jeu a été développé afin de permettre la validation des techniques utilisées pour *Space Invaders* sur un jeu plus complexe. Ainsi, *Lode Runner* fut développé en utilisant les abstractions de haut niveau que permettent les fonctions de première classe combinées avec un système d'objets très polyvalent et la conception de nouvelles formes spéciales. En utilisant, cette fois-ci, une approche conservatrice sur le contrôle de flot, le résultat fut très clair : le jeu a été développé rapidement et sans problèmes importants.

Malgré les problèmes reliés à l'ordonnancement des tâches dans *Space Invaders*, l'analyse des performances de ces jeux a permis de démontrer qu'ils correspondent bien aux besoins nécessaires pour le développement d'un jeu. L'utilisation de la mémoire est demeurée faible et les pauses dues à la récupération mémoire sont minimales. De plus, *Lode Runner* présente une utilisation du CPU relativement faible. Par contre, puisque ces jeux ont été originalement développés trente ans auparavant, ils ne peuvent qu'être grossièrement comparés aux jeux vidéo modernes en terme de complexité algorithmique et d'utilisation de la mémoire. Cette différence entre les jeux développés et des jeux mo-

derne rend une extrapolation des résultats obtenus difficile. Ainsi, ce mémoire présente un aperçu du potentiel de l'utilisation de Scheme dans le développement de jeu vidéo et, les conclusions tirées ici devront être validées dans le futur par le développement de jeux modernes en Scheme.

Il serait maintenant intéressant de voir quel serait l'impact de l'implantation en Scheme d'un jeu de qualité commerciale afin de tester plus profondément les limites que pourrait imposer le langage sur des jeux plus exigeants tant en utilisation processeur qu'en utilisation mémoire.

CHAPITRE 7

TRAVAUX RELIÉS

Les langages de programmation de la famille Lisp ont déjà été utilisés pour produire des jeux vidéo commerciaux de très bonne qualité. Certains parmi ceux-ci seront cités et une revue de l'expérience acquise par les développeurs de ces derniers sera exposée dans le but de comparer leurs expériences avec celles acquises pour le développement de *Space Invaders* et de *Lode Runner*.

Avant tout, nous comparerons les différents langages utilisés dans l'industrie du jeu vidéo avec Scheme afin d'en faire ressortir les forces et faiblesses. Les différences mènent à des méthodes de développement qui seront nécessairement divergentes. Ainsi, un portrait des techniques de programmation utilisées dans l'industrie est dressé dans le but de bien situer les travaux présentés dans ce mémoire.

7.1 Comparaison de langages

L'industrie du jeu vidéo est déjà bien établie et donc, certains langages sont favorisés quant au développement de jeux vidéo par les compagnies oeuvrant dans le domaine. Cette section présente ces langages dans le but de les comparer à Scheme afin de mettre en lumière la possibilité de les remplacer par ce dernier.

Les critères de comparaison utilisés dans cette section comprennent les fonctionnalités offertes par ces langages, leur typage, leur niveau d'abstraction ainsi que leurs modes d'utilisations.

7.1.1 C++

Le langage principalement utilisé pour le développement de jeux vidéo est sans aucun doute le langage C++ [6]. Ce dernier est une version orientée objet du langage C [5] qui est aussi très répandu et apprécié pour être un langage suffisamment de bas niveau pour permettre d'écrire du code très efficace. C++ apporte les paradigmes de

la programmation orientée objet en introduisant des classes à héritage multiple et des fonctions virtuelles permettant un *dispatch* dynamique simple. C++ possède également un système de gestion d'exceptions et de *templates*.

Ainsi, le langage C++ semble être grandement apprécié pour son mélange de concepts de haut niveau comme la programmation orientée objet et de bas niveau, tel que la gestion de mémoire manuelle. Il diffère grandement d'un langage de haut niveau tel que Scheme. En effet, il ne possède pas de fermetures, de continuations et de système de macros évolué. La programmation en C++ est dite impérative parce que le résultat du programme dépend intrinsèquement de l'état de celui-ci. De plus, puisqu'il s'agit d'un langage typé statiquement, le dépistage d'erreurs de programmation est plus ardu.

Il en découle donc que C++ offre beaucoup moins de concepts de haut niveau, réduisant ainsi la facilité d'abstraction et de programmation en général pour donner accès à de bonnes performances. À la lumière de ces différences, C++ semble un bon choix de langage pour des plate-formes possédant un matériel limité, tant en mémoire que puissance du processeur. Par contre, sur des machines plus puissantes, Scheme semble plus approprié parce qu'il apporte beaucoup plus de puissance d'abstraction aux programmeurs. Toutefois, une utilisation de C++ pour implanter les parties critiques du moteur d'un jeu et de Scheme pour développer la partie *gameplay* de ce dernier semble aussi une bonne approche sur des plate-formes évoluées afin d'optimiser les performances du jeu.

Un net avantage du langage C++ face à Scheme est la quantité astronomique de bibliothèques de tous genres disponibles aux programmeurs. Toutefois, l'interface aux fonctions étrangères de Gambit-C rend accessible l'utilisation de ces bibliothèques en Scheme avec peu d'efforts. Dans le cadre des travaux présentés dans le chapitre 6, l'utilisation de telles bibliothèques a été minimisée dans le but d'avoir la plus grande base de code Scheme possible afin de ne pas biaiser l'expérience faite. L'utilisation de bibliothèques C/C++ pour le développement d'applications commerciales représente certainement un net avantage. Par exemple, le chargement d'images pourrait être fait par l'une des nombreuses

librairies disponibles.

Aussi, C++ bénéficie de la présence d'une grande variété d'outils de développement tels des IDE (*Integrated Development Environment*) ou des systèmes avancés de profilages d'exécutions. Ces derniers facilitent grandement l'utilisation du langage C++ en permettant aux développeurs d'utiliser des débogueurs visuels avancés, des outils de refactorisation, des profileurs permettant de trouver les parties de code les plus utilisées (autant dans le processeur principal que sur la carte graphique), *etc.* Comme mentionné dans le chapitre 3, l'implantation Gambit-C du langage Scheme possèdent un débogueur dynamique très sophistiqué, mais elle ne possède pas une telle intégration dans un seul environnement de développement. Toutefois, puisque le système Gambit-C compile le langage Scheme vers le langage C ou C++, il serait possible d'utiliser les IDE disponibles pour C++ pour du code écrit en Scheme *dans une certaine mesure*. Il serait ainsi possible de, par exemple, visualiser les résultats de profilages fait sur la carte graphique, mais cette usage serait très peu adapté aux besoins du langage Scheme et viendrait briser tout l'intérêt d'un environnement de développement intégré en utilisant ces outils de manière externes au développement du code Scheme.

Il est également important de souligner le facteur de la main d'oeuvre pour l'adoption de Scheme pour remplacer C++. En effet, la quantité de programmeurs maîtrisant le langage Scheme est très restreinte en comparaison au nombre de programmeurs C++ et donc, même si Scheme pourrait se révéler un meilleur choix technologique pour le développement d'un jeu, il est possible que l'embauche de suffisamment de programmeurs Scheme soit trop difficile et coûteuse vis-à-vis l'embauche de programmeurs C++ beaucoup plus nombreux. Un tel phénomène pourrait ainsi rendre C++ plus attrayant et moins risqué aux yeux d'une entreprise.

7.1.2 Lua

Le langage Lua [56] est fréquemment utilisé non pas pour implanter le moteur d'un jeu vidéo, mais plutôt pour effectuer le *scripting* de celui-ci, *i.e.* le développement

dynamique de niveaux, d'environnements, *etc.* Lua a été utilisé notamment dans le jeu *World Of Warcraft* afin de permettre aux joueurs d'ajouter des éléments au jeu de manière dynamique.

Ce langage est reconnu pour avoir un environnement d'exécution très léger, environ 150 kilo-octets, ce qui fait en sorte que l'ajout de ce système a un faible coût en espace mémoire, ce qui est désirable lorsque cette intégration est faite sur des plate-formes au matériel plus limité comme des consoles portables. Lua contient certains éléments de la programmation dite de haut niveau dont les fermetures, une gestion de mémoire automatique, l'optimisation d'appels terminaux et un système de coroutines. Il fournit aussi un système objet limité à la Javascript [24], basé sur le prototypage d'objets.

Ainsi, malgré le fait que ce dernier contient plusieurs aspects de la programmation de haut niveau, Lua possède plusieurs différences au langage Scheme. Il ne supporte pas la réification de continuations et ne possède pas de système de macros. Ainsi, l'absence de ces fonctionnalités rend l'abstraction et la modularité, utiles pour de gros projets, moins accessibles. Il en résulte que Lua est moins approprié pour écrire le moteur de jeux vidéo non triviaux. Toutefois, puisque l'interprète du langage est très rapide, il pourrait être envisagé pour le développement de jeux à plus petite échelle.

Plusieurs environnements de développements intégrés sont disponibles pour Lua, mais ceux-ci demeurent très rudimentaires et supportent des fonctionnalités de base tels des déboggeurs visuels.

Pour les mêmes raisons que Lua, d'autres langages de programmations sont également utilisés afin d'effectuer le *scripting* de jeux. Parmi ceux-ci, on retrouve principalement Python [57] et Javascript [24] qui sont très similaires à Lua et donc constituent eux aussi un sous ensemble de Scheme.

7.2 Sommaire

Ainsi, les langages C++ et Lua offrent des fonctionnalités différentes de celles offertes en Scheme. Le tableau 7.1 effectue un résumé des différences entre ces langages.

Points de comparaison	C++	Lua	Scheme
Typage	statique	dynamique	dynamique
Gestion mémoire	manuelle	automatique	automatique
Fermetures	non	oui	oui
Macros	textuelles	aucune	procédurales
Réification de continuations	non	non	oui
Mode usuel d'utilisation	compilé	interprété	compilé et interprété
Dépistage d'erreurs	bas niveau	bas niveau	haut niveau
IDE	très développé	de base	inexistant

TABLEAU 7.1 – Comparaison des langages C++, Lua et Scheme

Le langage C++ offre de très bonnes performances grâce à son modèle d'exécution près de l'architecture de la plate-forme utilisée. Puisqu'il n'offre que très peu de fonctionnalités de haut niveau, il constitue un choix idéal pour l'implantation de jeux sur des plate-formes aux ressources très limitées, tel que c'est le cas sur les consoles portables ou sur les téléphones cellulaires.

Le langage Lua, quant à lui, est un sous-ensemble du langage Scheme. La présence de gestion de mémoire automatique, de fermeture et d'un typage dynamique font de lui un très bon candidat comme langage de script afin d'évaluer dynamiquement des parties expérimentales du jeu dans le but de raffiner le *gameplay*.

Scheme offre beaucoup de fonctionnalités de haut niveau telles que des fermetures, de la réification de continuations ou encore des macros procédurales. Son coût plus élevé en allocation mémoire fait de lui un excellent choix pour le développement de moteurs de jeu sur des plate-formes modernes ou comme engin de *scripting* sur des plate-formes plus limitées sur la mémoire.

Le langage C++ offre à ses programmeurs la possibilité d'utiliser des environnements de développement intégrés avancés qui donnent un accès direct à des outils de développement et d'optimisation. Les IDE disponibles pour Lua possèdent un sous-ensemble des fonctionnalités de ceux offerts pour C++, mais permettent un accès facile aux fonctionnalités de bases telles un débogueur visuel. Les développeurs Gambit-C ne possède pas de tels outils de développement intégrés, mais peuvent utiliser de manière indirecte les outils de développement offert pour les langages C et C++.

7.3 Jeux en Lisp

Malgré la domination du langage C++ dans l'industrie du jeu vidéo, certains jeux ont déjà été reconnus pour utiliser Lisp ou Scheme à différents degrés dans leur développement. Ceux-ci sont mis en lumière dans cette section afin de permettre de comparer l'expérience acquise par ces développeurs à celles acquises dans notre travail.

7.3.1 Naughty Dog

La compagnie Naughty Dog [58] est bien connue pour utiliser des dialectes de Lisp dans le développement de jeux vidéo. En effet, cette compagnie a écrit plusieurs jeux sur le *PlayStation 2* en utilisant le compilateur *GOAL/Game Oriented Assembly Lisp*, un système maison uniquement compilé vers du code machine pour le *PlayStation 2* qui comprenait notamment un système objet et des coroutines. Ainsi, le compilateur *GOAL* fut utilisé pour développer entièrement de nombreux jeux sur la console dont les jeux des séries *Crash Bandicoot* et *Jak and Dexter*.

Scott Shumaker, programmeur chef chez Naughty Dog, a commenté leur utilisation de *GOAL* dans un article sur le site web consacré au développement de jeux vidéo Gamasutra [59]. Dans ce dernier, il mentionne que le développement de leur propre compilateur de Lisp leur a permis d'utiliser plusieurs techniques qui n'auraient pas pu être envisagées avec d'autres langages. Entre autres, leur système, malgré le fait qu'il ne comprenait pas d'interprète, permettait la compilation de code et le chargement de celui-ci directement dans le jeu qui était en cours d'exécution sur une console *PlayStation 2*. Ils obtenaient ainsi de la programmation très dynamique basée sur du prototypage. Il mentionne également que *GOAL* comprenait un système de coroutines. Un exemple de code simplifié pour *GOAL* est donné dans la figure 7.1. Le code présenté a été simplifié dans le but d'illustrer l'utilisation des coroutines de ce jeu. On constate que le code conserve l'état de la variable `ii` qui se trouve à être, dans cet exemple, un compteur d'images rendues pour une animation nommée `idle`. L'utilisation de coroutines leur a permis de conserver l'état du calcul en cours de manière triviale.


```
(dotimes (ii (num-frames idle))
  (set! frame-num ii)
  (suspend))
```

FIGURE 7.1 – Exemple de code en GOAL utilisant le système de coroutines

Il mentionne également dans cet article que l'utilisation de leur système GOAL a été la source de plusieurs problèmes principalement dû au fait que GOAL fut développé par un seul programmeur Lisp dont lui seul en comprenait parfaitement l'étendue. Il mentionne aussi qu'il fut difficile de trouver de la main d'oeuvre maîtrisant le langage Lisp et que même ceux-ci devaient tout de même s'adapter au système utilisé.

Ils ont par la suite créé les jeux *Uncharted : Drake's fortune* et *Uncharted 2 : Among Thieves* qui sont rapidement devenus très populaires pour leurs scènes d'actions et leurs rendus graphiques exceptionnels. Dan Liebgold de Naughty Dog a donné une conférence sur le développement du jeu *Uncharted : Drake's fortune* [1] et mentionna qu'ils ont utilisé le langage Scheme afin d'exprimer les données du jeu de manière dynamique, mais que le moteur du jeu est écrit en C++. Ils profitent ainsi des performances qu'offrent C++ pour les zones critiques du jeu et utilisent un langage de plus haut niveau pour la description des données. Ils utilisent ainsi un schéma de compilation dynamique similaire à celui utilisé par GOAL. La figure 7.2 illustre la construction de données utilisées dans ce jeu. Cet exemple de code illustre bien le type d'utilisation faite du langage Scheme, *i.e.* que Scheme est utilisé non seulement comme langage de script, mais aussi pour la définition des données du langage.

```

(deftype vec4 (:align 16)
  ((x float) (y float)
   (z float) (w float :default 0)))

(deftype quaternion (:parent vec4) ())

(define (axis-angle->quat axis angle)
  (let ((sin-angle/2 (sin (* 0.5 angle))))
    (new quaternion
      :x (* (-> axis x) sin-angle/2)
      :y (* (-> axis y) sin-angle/2)
      :z (* (-> axis z) sin-angle/2)
      :w (cos (* 0.5 angle)))))

```

FIGURE 7.2 – Déclaration de données en Scheme utilisées dans le jeu *Uncharted : Drake's fortune* [1]

7.3.2 *QuantZ*

Récemment, le jeu *QuantZ* [60] fut développé presque exclusivement en utilisant le langage Scheme et le compilateur Gambit-C. Ce dernier constitue un jeu de casse-tête 3D similaire à *Tetris* où le but est d'agencer des billes de même couleur sur un cube. Les informations relatives au développement de ce jeu ont été acquises par l'auteur qui est toujours employé par le studio de développement de jeux vidéo *Gamerizon inc.*

L'utilisation de Scheme dans *QuantZ* a apporté plusieurs avantages technologiques. En effet, le dynamisme du langage a su profiter aux développeurs en leur permettant de débogger le jeu lors de son exécution. Cet aspect du langage permet la résolution rapide de problèmes et donc d'accélérer le développement du jeu.

Aussi, la syntaxe sous formes de S-expressions des programmes Scheme a été utilisée à profit en créant un système d'analyse de code permettant d'extraire toutes les chaînes de caractères présentes dans le jeu afin de créer des outils d'internationalisation. Ces outils permettent notamment la création automatique de fontes optimisées par langue qui contiennent uniquement les caractères utilisées dans chacune. L'utilisation de la fonction Scheme `read` combinée à la récursivité du langage ont facilement permis d'écrire un analyseur de code. Cette tâche aurait été beaucoup plus complexe dans un autre langage de programmation.

Les fermetures disponibles en Scheme ont été utilisées comme base d'un système d'interface graphique complet permettant la création de fenêtres, de boutons, *etc.* En utilisant ces fermetures pour conserver l'état de ces entités, on obtient ainsi un résultat similaire à l'utilisation d'un système de programmation orientée objet, où l'état est conservé dans ces fermetures, mais ne fournissant pas de fonctions génériques.

Afin d'éviter des problèmes de pauses aléatoires dues à la gestion mémoire automatique, plusieurs techniques ont été utilisées dans le développement de ce jeu. Dans un premier temps, le gestionnaire de mémoire est appelé explicitement avant le rendu de chaque image afin d'éviter d'accumuler trop de données mortes dans la mémoire et causer ainsi de longues pauses durant l'exécution du jeu. Aussi, la sérialisation de données disponible dans Gambit-C a été utilisée afin de transformer de grosses structures de données statiques (préférences, pointage de joueurs, ...) afin que celles-ci ne soient pas traversées par le gestionnaire avant le rendu de chaque image du jeu.

L'état du jeu a été implanté en utilisant un mécanisme de programmation fonctionnelle réactive basée sur Scheme permettant l'écriture élégante de niveaux grâce à la propagation de signaux correspondant à l'état du jeu. Aussi, le moteur de physique du jeu, écrit en C, a pu être intégré sans problèmes au reste du moteur de *QuantZ* en utilisant l'interface aux fonctions étrangères de Gambit-C.

Ainsi, le jeu *QuantZ* représente un exemple concret d'un jeu commercial ayant utilisé Scheme comme langage principal de développement. L'utilisation faite du langage a su tirer profit de la plupart des particularité de Scheme dont notamment l'utilisation de fonctions de première classe et du dynamisme du système Gambit-C.

7.4 Autres version des jeux développés

Cette section présente les versions originales des jeux développés dans le but de les comparer aux versions écrites pour ce mémoire. Puisque le code source des versions originales n'est pas disponible, une comparaison algorithmique ne peut pas être effectuée. Ainsi, une approche plus pratique est utilisée en effectuant une comparaison

de performance des différentes versions de *Space Invaders* et de *Lode Runner*.

7.4.1 *Space Invaders*

Tel que mentionné dans le chapitre 2, le jeu original de *Space Invaders* date de 1978 sur console arcade. Le concepteur avait dû, à l'époque, non seulement concevoir le jeu, mais aussi le matériel sur lequel le jeu fonctionnait. À lui seul, le développement complet du jeu lui prit environ un an. Le processeur utilisé était un 8080 d'Intel offrant une vitesse d'horloge allant jusqu'à 2 MHz. Le tableau 7.2 illustre ces différences avec le jeu développé pour ce mémoire. On constate que l'effort de production du jeu original était plus grand, mais en tenant compte du fait que le développement initial comprenait aussi la mise au point du matériel utilisé, on peut estimer des temps de développement équivalents pour les deux versions. Toutefois, le code original, écrit en langage assembleur est certainement très spécifique au jeu écrit. Par contre, le code écrit pour la version utilisée dans ce mémoire apporte des bibliothèques qui permettent de développer d'autres jeux vidéo plus facilement. D'autre part, un facteur de mille entre la vitesse d'horloge de l'architecture originale face aux architectures modernes des processeurs est observé, pouvant laisser croire que notre implantation est moins efficace de ce facteur.

En ce qui concerne le jeu *Lode Runner*, la version finale du jeu fut publiée en 1983 sur l'ordinateur *Apple II+* et écrit en langage assembleur. La durée de développement de ce jeu est difficile à donner, puisque le jeu final est le fruit de la troisième réécriture du code source. De plus, ce concepteur effectuait à temps partiel, puis à temps plein le développement du jeu. On peut toutefois estimer le développement total à environ un an, en incluant la conception des 150 niveaux du jeu et l'écriture de toutes les versions du code. La vitesse d'horloge du processeur utilisé était similairement de 1 MHz.

En assumant que les jeux originaux utilisaient 30 images par secondes pour leur rendu graphique et en tenant compte des performances des jeux développés présentées dans le tableau 6.1, on obtient un ratio indiquant que la version Scheme du jeu est environ 350 fois plus lente que la version originale du jeu. En effet, l'utilisation pro-

cesseur de nos jeux sont de 40 % pour *Space Invaders* et de 30 % pour *Lode Runner*. Cette différence de performance entre les jeux originaux écrits il y a trente ans et les réécriture faites pour ce mémoire peut être expliquée par plusieurs facteurs. D’une part, les jeux originaux sont écrits directement en langage assembleur et donc, beaucoup d’optimisations de très bas niveaux ont pu être faites grâce à l’accès presque direct aux ressources du matériel, ce qui n’est pas possible pour des jeux écrits en Scheme et qui sont exécutés dans un système d’exploitation complexe exécutant des dizaines de processus de manière concurrente. D’autre part, comme il a déjà été mentionné, les jeux écrits pour la rédaction de ce document possèdent de puissantes abstractions tels un système de programmation orientée objet ou un système de coroutines. Ces dernières imposent un coût de performances tel que mentionné dans les tableaux 4.3 et 5.1. Les abstractions du langage Scheme apportent elles aussi un certain coût de performance en comparaison à du code écrit directement en langage assembleur.

Les temps de développement des versions réécrite des jeux sont comparable aux temps requis pour le développement des premières versions officielles de ces derniers. Toutefois, le temps de développements de six mois pris pour la réécriture comprend l’élaboration de plusieurs librairies facilitant le développement ultérieur de nouveaux jeux, ce qui n’est probablement pas le cas des versions originales écrites en langage de bas niveau.

Titre	Année de sortie	Vitesse CPU	Durée de développement
Space Invaders (original)	1978	2 MHz	1 année*homme
Space Invaders (mémoire)	2008	1.6 GHz	1/2 année*homme
Lode Runner (original)	1983	1 MHz	1/2 année*homme
Lode Runner (mémoire)	2009	1.6 GHz	1 année*homme

TABLEAU 7.2 – Comparaison entre les versions originales et réécrites des jeux développés.

Ainsi, la réécriture d’anciens jeux en utilisant des langages de haut niveaux implique un coût de performance qui est compensé par un développement qui semble non seulement plus rapide, mais qui favorise la réutilisation de code pour le développement ultérieurs de nouveaux jeux.

7.5 Conclusion

Plusieurs langages sont utilisés dans l'industrie du jeu vidéo dont, notamment les langages Lua et C++. Après une comparaison au langage Scheme, on constate que ces derniers répondent à des besoins précis présents dans les jeux vidéo. Lua propose un développement plus dynamique tandis que C++ permet d'obtenir de très bonnes performances au coût d'avoir moins de puissance d'abstraction. Le choix du langage le plus approprié pour le développement d'un jeu devrait donc être fait en prenant en considération le matériel utilisé afin d'obtenir le plus de puissance expressive que possible, sans toutefois avoir un coût relié à ces abstractions trop élevées.

La présence d'environnement de développement intégrées rend attrayant l'utilisation des langages C++ et Lua, mais la plupart des fonctionnalités offertes peuvent être utilisées pour le langage Scheme de manière indirecte.

Aussi, Lisp et Scheme ont déjà été utilisés dans l'industrie du jeu vidéo pour développer des jeux de grande qualité tels *Jak and Dexter*, *Uncharted : Drake's Fortune* et *QuantZ*. L'expérience acquise par les développeurs de ces jeux est très positive. On constate de manière unanime ces langages leur ont fourni un environnement de développement dynamique. L'extensibilité du langage a aussi joué à l'avantage de ceux-ci. Ainsi, malgré la faible adoption de Scheme pour l'écriture de jeux vidéo commerciaux par les compagnies de l'industrie, les équipes l'ayant utilisé semblent l'avoir beaucoup apprécié et surtout, l'utilisent toujours dans le développement de leurs jeux.

Finalement, les versions réécrites de *Space Invaders* et de *Lode Runner* lorsque comparées aux versions originales semblent démontrer une utilisation du processeur approximativement 350 fois plus exigeante, mais utilisent des abstractions de haut niveau permettant la réutilisation du code écrit pour l'écriture de nouveaux jeux, contrairement aux versions originales écrites en langage assembleur.

CHAPITRE 8

CONCLUSION

L'expérience acquise lors du développement des jeux *Space Invaders* et *Lode Runner* permet de faire le point sur les forces et les faiblesses de l'utilisation d'un langage de programmation fonctionnelle et dynamique tel que Scheme pour le développement de jeux vidéo.

L'expressivité du langage Scheme, notamment grâce aux fonctions de première classe et aux macros procédurales, a permis de développer un système de programmation orientée objet et de l'intégrer directement au langage. Aussi, la réification de continuations a rendu possible le développement d'un système de coroutines élaboré. L'utilisation de ces systèmes dans la programmation de jeux vidéo fut bénéfique quant à l'amélioration du code produit grâce à l'introduction d'une bonne modularité fournie par les fonctions génériques du système objet et une grande simplicité pour l'implantation de parties multi-joueurs grâce à l'utilisation de coroutines pour encapsuler les parties se déroulant en parallèle. Le coût de développement de tels modules, en utilisant des langages donnant accès à moins de puissance d'abstraction, aurait certainement rendu le développement de ceux-ci plus difficile. Par conséquent, Scheme s'est révélé un langage puissant pour le développement de ces outils indispensables.

Le dynamisme du langage a également joué un rôle implicite important pour le développement de ces jeux. En effet, la présence d'un débogueur dynamique et efficace a apporté une grande flexibilité au développement et permis de développer efficacement ces deux jeux. En effet, malgré le fait que les erreurs de programmation sont signalées durant l'exécution du jeu, la résolution de celles-ci se font rapidement grâce à la possibilité d'introspection de l'état du programme et d'évaluation dynamique de code.

Aussi, le langage Scheme a permis d'expérimenter sur le contrôle de flot des jeux en permettant de développer des systèmes non-triviaux de simulation à événements

discrets et de simulation par agents (coroutines). La modification du flot de contrôle du premier jeu a pu être fait sans trop d'efforts grâce aux abstractions faites, entre autres, par le système objet.

Toutefois, la modularité de programmes apporte généralement des coûts en performance pour ceux-ci. Ainsi, une difficulté d'utiliser un langage comme Scheme face à la programmation de jeux vidéo réside entre trouver le bon équilibre entre le niveau d'abstraction et la spécialisation de code (optimisations). Cette balance varie grandement en fonction de l'architecture du matériel sur lequel le jeu sera déployé. S'il s'agit d'une plate-forme au matériel très limité, alors le coût d'abstraction pourrait s'avérer trop grand. Dans un tel cas, l'utilisation de langages de plus bas niveau, comme C++, semble plus appropriée afin de bâtir le moteur du jeu, mais Scheme pourrait alors être utilisé comme langage de script afin de développer dynamiquement le *gameplay* du jeu. Par contre, s'il s'agit de plate-formes performantes, comme c'est le cas pour les jeux développés dans le cadre de ce mémoire, alors la possibilité de modulariser le code devient plus accessible et les performances deviennent moins critiques, face à la réutilisabilité potentielle du code écrit. Scheme devient alors un candidat pourrait être idéal.

Un problème potentiel qui n'a pas été adressé dans ce mémoire est relié à l'allocation et la gestion de mémoire automatique utilisée par le langage Scheme. En effet, puisque Scheme est un langage fonctionnel, ce dernier effectue généralement beaucoup d'allocations mémoire. Ainsi, il est possible que sur des systèmes où le coût d'allocation mémoire est élevé, l'utilisation d'un tel langage devienne trop coûteux. De même, la présence du gestionnaire automatique de mémoire pourrait potentiellement engendrer des pauses de durées aléatoires du jeu qui sont indésirables. Ces phénomènes n'ont pas été observés dans les jeux développés pour ce mémoire puisqu'ils demeureraient trop simples. Des jeux plus modernes possèdent généralement un plus grand nombre d'entités de manière simultanée durant une partie, augmentant ainsi le stress sur la mémoire et donc, sur la récupération automatique de celle-ci.

Il serait donc intéressant d'utiliser les techniques présentées dans ce mémoire afin de développer des jeux de plus grande envergure dans le but d'obtenir une meilleure idée sur l'impacte d'une utilisation plus intensive de la mémoire dans un jeu. Un tel jeu écrit en Scheme permettrait aussi de pousser plus loin l'aspect dynamique du langage par l'écriture d'outils d'écriture de scripts qui pourraient être utilisés en temps réel lors de la conception des niveaux. L'utilisation des techniques présentées dans ce mémoire permettraient ainsi une utilisation complète du langage et donc de fournir un aperçu global des performances qu'offre Scheme pour le développement de jeux vidéo.

BIBLIOGRAPHIE

- [1] Dan Liebgold. Adventures in Data Compilation : Uncharted : Drake's Fortune. Game Developers Conference, 2008.
- [2] NPD Group. 2007 U.S. Video Game And PC Game Sales Exceed \$18.8 Billion Marking Third Consecutive Year Of Record-Breaking Sales.
http://www.npd.com/press/releases/press_080131b.html, 2008.
- [3] NPD Group. U.S. Retail Sales of Non-Games Software Experience near Double-Digit Decline in 2008.
http://www.npd.com/press/releases/press_090217.html, 2009.
- [4] Chris Morris. Special to CNBC.com — 12 Jun 2009 — 05 :12 PM ET.
<http://www.cnbc.com/id/31331241>, Sep 2009.
- [5] Dennis M. Ritchie. The development of the C language. In *HOPL-II : The second ACM SIGPLAN conference on History of programming languages*, pages 201–208, New York, NY, USA, 1993. ACM.
- [6] Margaret A. Ellis and Bjarne Stroustrup. *The annotated C++ reference manual*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1990.
- [7] ECMA. C# Language Specification.
<http://www.ecma-international.org/publications/standards/Ecma-334.htm>, 2006.
- [8] Richard Kelsey, William Clinger, and Jonathan Rees (Editors). Revised⁵ report on the algorithmic language Scheme. *ACM SIGPLAN Notices*, 33(9) :26–76, 1998.
- [9] Marc Feeley. Gambit-C. <http://www.iro.umontreal.ca/~gambit>.
- [10] Marc Feeley. Gambit-C vs. Bigloo vs. Chicken vs. MzScheme vs. Scheme48.
<http://www.iro.umontreal.ca/~gambit/bench.html>, Sep 2009.
- [11] Leonard Herman, Jer Horwitz, Steve Kent, Skyler Miller. The History Of Video Games.
<http://www.gamespot.com/gamespot/features/video/hov/index.html>, Sep 2009.
- [12] Johnny L. Wilson Rusel DeMaria. *High score ! : the illustrated history of electronic games*. McGraw Hill, 2004.
- [13] Thomas T. Goldsmith Jr. Cathode Ray Tube Amusement Device. U.S. Patent #2455992, 1948.
- [14] Dennis Ritchie and Ken Thompson. Unix Past.
http://www.unix.org/what_is_unix/history_timeline.html, Sep 2009.
- [15] Joshua Walrath. 30 Frames per Second vs. 60 Frames per Second.
http://www.daniele.ch/school/30vs60/30vs60_1.html, Sep 1999.

- [16] Erik Meijer and Peter Drayton. Static typing where possible, dynamic typing when needed : The end of the cold war between programming languages. 2005.
- [17] Harold Abelson, Gerald Jay Sussman, and Julie Sussman. *Structure and interpretation of computer programs*. MIT Press, Cambridge, Mass., 1996.
- [18] John McCarthy. Recursive functions of symbolic expressions and their computation by machine, Part I. *Commun. ACM*, 3(4) :184–195, 1960.
- [19] Guy L. Steele. *Common Lisp the Language*. Digital Press, 2nd edition edition, 1984.
- [20] Gerald Jay Sussman. Scheme : An interpreter for extended lambda calculus. In *Memo 349, MIT AI Lab*, 1975.
- [21] Carl Hewitt, Peter Bishop, and Richard Steiger. A Universal Modular ACTOR Formalism for Artificial Intelligence. 1973.
- [22] Gerald Jay Sussman and Guy L. Steele, Jr. The First Report on Scheme Revisited. *Higher Order Symbol. Comput.*, 11(4) :399–404, 1998.
- [23] Scheme Request For Implementation.
<http://srfi.schemers.org/>, Sep 2009.
- [24] E. C. M. A. International. *ECMA-262 : ECMAScript Language Specification*. ECMA (European Association for Standardizing Information and Communication Systems), Geneva, Switzerland, third edition, December 1999.
- [25] Ruby Programming Language.
<http://www.ruby-lang.org/>, Sep 2009.
- [26] Perl Language.
<http://www.perl.org/>, Sep 2009.
- [27] Alonzo Church. An Unsolvable Problem of Elementary Number Theory. *American Journal of Mathematics*, 58(2) :345–363, 1936.
- [28] Robin Milner, Mads Tofte, and Robert Harper. *The definition of Standard ML*. MIT Press, Cambridge, MA, USA, 1990.
- [29] Simon Peyton Jones. *Haskell 98 language and libraries : the revised report*. Cambridge University Press, Cambridge, MA, USA, 2003.
- [30] Eugene E Kohlbecker. *Syntactic extensions in the programming language LISP*. PhD thesis, Bloomington, IN, USA, 1986.
- [31] Paul R. Wilson. Uniprocessor Garbage Collection Techniques. In *IWMM '92 : Proceedings of the International Workshop on Memory Management*, pages 1–42, London, UK, 1992. Springer-Verlag.
- [32] David Plainfossé and Marc Shapiro. A Survey of Distributed Garbage Collection Techniques. In *IWMM '95 : Proceedings of the International Workshop on Memory Management*, pages 211–249, London, UK, 1995. Springer-Verlag.

- [33] Robert R. Fenichel and Jerome C. Yochelson. A LISP garbage-collector for virtual-memory computer systems. *Commun. ACM*, 12(11) :611–612, 1969.
- [34] Henry Lieberman and Carl Hewitt. A real-time garbage collector based on the lifetimes of objects. *Commun. ACM*, 26(6) :419–429, 1983.
- [35] Henry C. Baker, Jr. and Carl Hewitt. The incremental garbage collection of processes. In *Proceedings of the 1977 symposium on Artificial intelligence and programming languages*, pages 55–59, New York, NY, USA, 1977. ACM.
- [36] Edsger W. Dijkstra, Leslie Lamport, A. J. Martin, C. S. Scholten, and E. F. M. Steffens. On-the-fly garbage collection : an exercise in cooperation. *Commun. ACM*, 21(11) :966–975, 1978.
- [37] Benjamin Zorn. Barrier Methods for Garbage Collection. Technical report, 1990.
- [38] Marc Feeley. A Better API for First-Class Continuations, 2001.
- [39] Richard Kelsey. Scheme Request For Implementation 9 : Defining Record Types. <http://srfi.schemers.org/srfi-9/srfi-9.html>, Sep 2009.
- [40] James Gosling, Bill Joy, Guy Steele, and Gilad Bracha. *The Java Language Specification, Third Edition*. Addison-Wesley Longman, Amsterdam, 3 edition, June 2005.
- [41] Daniel G. Bobrow, Linda G. DeMichiel, Richard P. Gabriel, Sonya E. Keene, Gregor Kiczales, and David A. Moon. Common Lisp Object System Specification X2J13 Document 88-002R. *SIGPLAN Notices*, 23(Special Issue) :1.1–2.94, 1988.
- [42] Gregor Kiczales, Jim D. Rivieres, and Daniel G. Bobrow. *The Art of the Metaobject Protocol*. The MIT Press, July 1991.
- [43] Christian Queinnec and Ecole Polytechnique. Meroon V3 : A Small, Efficient and Enhanced Object System. Technical report, Ecole Polytechnique and INRIA-Rocquencourt, 1993.
- [44] Kenneth Dickey. Thinking Scheme. Workshop on Scheme and Functional Programming., september 2008.
- [45] Marc Feeley. SRFI 18 : Multithreading support. <http://srfi.schemers.org/srfi-18/srfi-18.html>.
- [46] Marc Feeley Guillaume Germain and Stefan Monnier. Termite : a Lisp for Distributed Computing. 2nd European LISP and Scheme Workshop. <http://lisp-ecoop05.bknr.net/submission/19654.html>, July 2005.
- [47] Ericsson Computer Science Laboratory. Erlang/otp. <http://www.erlang.org>.
- [48] Guillaume Germain. Sources du système termite. <http://toute.ca/termite>.
- [49] Zhanyong Wan and Paul Hudak. Functional reactive programming from first principles. *SIGPLAN Not.*, 35(5) :242–252, 2000.

- [50] Edsger W. Dijkstra. The structure of the “the”-multiprogramming system. In *SOSP '67 : Proceedings of the first ACM symposium on Operating System Principles*, pages 10.1–10.6, New York, NY, USA, 1967. ACM.
- [51] Marc Feeley and Martin Larose. Compiling Erlang to Scheme. In *PLILP '98/ALP '98 : Proceedings of the 10th International Symposium on Principles of Declarative Programming*, pages 300–317, London, UK, 1998. Springer-Verlag.
- [52] Antony Courtney, Henrik Nilsson, and John Peterson. The Yampa arcade. In *Haskell '03 : Proceedings of the 2003 ACM SIGPLAN workshop on Haskell*, pages 7–18, New York, NY, USA, 2003. ACM.
- [53] Multiple Arcade Machine Emulator (MAME).
<http://mamedev.org/>, Sep 2009.
- [54] Simple Directmedia Layer (SDL).
<http://www.libsdl.org/>, Sep 2009.
- [55] Open Graphic Library (OpenGL).
<http://www.opengl.org/>, Sep 2009.
- [56] Langage de programmation Lua.
<http://www.lua.org/>, Sep 2009.
- [57] Langage de programmation Python.
<http://docs.python.org/reference/>, Sep 2009.
- [58] Naughty Dog.
<http://www.naughtydog.com/>, Sep 2009.
- [59] Postmortem : Naughty Dog’s Jak and Daxter : the Precursor Legacy.
http://www.gamasutra.com/view/feature/2985/postmortem_naughty_dogs_jak_and_.php, Sep 2009.
- [60] QuantZ.
<http://www.quantzgame.com/>, Sep 2009.