HAL
open science

Couplage de systèmes dynamiques pour l'émergence de comportement en environnement virtuel : application au rebond de balle

Anne Jeannin-Girardon

▶ To cite this version:

Anne Jeannin-Girardon. Couplage de systèmes dynamiques pour l'émergence de comportement en environnement virtuel : application au rebond de balle. Synthèse d'image et réalité virtuelle [cs.GR]. 2011. dumas-00636431

HAL Id: dumas-00636431
https://dumas.ccsd.cnrs.fr/dumas-00636431v1

Submitted on 27 Oct 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Rapport de stage

Université de Bretagne Occidentale

Master 2 Recherche en informatique

Couplage de systèmes dynamiques pour l'émergence de comportement en environnement virtuel : application au rebond de balle

2 juin 2011

LISyC
Équipe ARéVi

Réalisé par :
Anne Jeannin-Girardon

Encadrants :
Pierre De Loor
Cédric Buche

Résumé

Ce document aborde les notions d'autonomie et d'interactivité pour des agents en environnement virtuel. L'utilisation d'approches à base de systèmes dynamiques couplés issues de la psychologie cognitive écologique doit permettre l'émergence de comportements adaptatifs.
L'utilisation de ce type d'approche vise à contourner les limitations des approches existantes et à l'obtention d'un meilleur réalisme dans la dynamique de l'interaction d'agents autonomes.

Mots clefs : agents interactifs, environnements virtuels, systèmes dynamiques, couplage, émergence de comportement, perception et action.

Table des matières

Résumé **i**

Introduction **1**

1 État de l'art **2**
1.1 Contexte 2
1.2 Agents interactifs temps réel 4
1.2.1 Interaction et adaptation chez les humanoïdes virtuels 4
1.2.2 Émergence de synchronicités chez des robots 7
1.3 Dynamique du comportement en psychologie cognitive 8
1.3.1 Introduction aux systèmes dynamiques 9
1.3.2 Dynamique comportementale (Warren, 2006) 9
1.3.3 Limitations de la dynamique comportementale 11
1.4 Synthèse 13

2 Cas d'étude : le rebond de balle **14**
2.1 Généralités 14
2.2 Modes de contrôle 15
2.2.1 Contrôle actif de la raquette 16
2.3 Le système balle/raquette 17

3 Modélisation du système balle/raquette **19**
3.1 Environnement de simulation 19
3.2 Oscillateur et pilotage de la raquette 19
3.3 Algorithme de contrôle 21
3.3.1 Modulation de période 21
3.3.2 Modulation d'amplitude 21

4 Résultats & Discussion **24**
4.1 Rebond passif 24
4.1.1 Méthode 24
4.1.2 Conditions environnementales constantes 24
4.1.3 Conditions environnementales variables au cours d'une même simulation . 27
4.1.4 Synthèse 29
4.2 Rebond avec contrôle actif 30
4.2.1 Méthode 30
4.2.2 Caractérisation 30
4.2.3 Conditions environnementales constantes au cours d'une même simulation 32
4.2.4 Conditions environnementales variables au cours d'une même simulation . 35
4.2.5 Synthèse 37

Conclusion et perspectives **38**

Références **39**

Introduction

Les travaux réalisés au Centre Européen de Réalité Virtuelle (CERV) abordent la notion de l'autonomie dans les systèmes de réalité virtuelle (immersive ou non). Dans le cadre d'interactions entre agents, ou mieux, avec des humains, la problématique soulevée est de savoir comment coupler les notions d'autonomie et d'interactivité. L'utilisation d'approches dynamiques vise à l'obtention de plus de réalisme dans la *dynamique de l'interaction*.

Nous définissons la dynamique de l'interaction par *ce qui caractérise l'évolution des échanges entre deux acteurs participant à une activité commune*. Ce type d'interaction fait apparaître des accélérations, des ralentissements, et on observe l'émergence de comportements tels que le turn-taking, l'imitation ou encore la synchronisation.

L'objet de ce travail est d'évaluer la possibilité d'utiliser des approches à base de systèmes dynamiques issus de la psychologie cognitive écologique afin de doter des entités autonomes d'un meilleur réalisme dans la dynamique de leurs interactions. L'idée de ce type d'approche est d'effectuer un couplage entre les systèmes dynamiques en jeu afin de donner lieu à une émergence de comportement, permettant l'obtention de comportements modulables et adaptatifs.

L'utilisation d'une telle approche doit à terme permettre de renforcer les interactions entre agents en environnement virtuel. Dans un premier temps toutefois, il est nécessaire de mettre en œuvre une étude de cas n'implicant pas nécessairement plusieurs agents afin d'évaluer les résultats obtenus par cette approche. Le cas d'étude que nous avons traité dans le cas présent est une tâche consistant à faire rebondir une balle sur une raquette ; on se place donc dans un cadre d'interactions agent/environnement.

La partie 1 dresse un état de l'art des agents interactifs virtuels ou robotisés. Les limitations d' approches à base de règles ou de raisonnement mènent à l'introduction de la dynamique du comportement telle qu'étudiée en psychologie cognitive écologique à travers l'approche de Warren (2006). Son approche, basée sur le couplage de systèmes dynamiques autonomes doit donner lieu à l'observation d'émergence de comportement telle que nous l'avons définie à travers la dynamique de l'interaction.

La partie 2 présente une étude de cas permettant la mise en œuvre de l'approche de Warren (2006) : La tâche du cas d'étude consiste à faire rebondir une balle sur une raquettede manière rythmique . Le choix d'un cas d'étude est indispensable pour évaluer la possibilité de faire émerger des comportements par des systèmes dynamiques puisqu'il faut déterminer les informations nécessaires au couplage des systèmes utilisés dans la tâche.

La partie 3 concerne la modélisation de cette tâche en environnement virtuel. Un premier point concerne la descriptions des systèmes dynamiques utilisés pour la tâche et un second point traite du couplage de ces systèmes dans l'environnement.

La partie 4 présente les résultats obtenus en simulation en environnement virtuel et propose une réflexion sur les résultats par rapport à la problématique du sujet.

Partie 1

État de l'art

1.1 Contexte

Les humanoïdes peuplant les environnements virtuels bénéficient aujourd'hui d'animations gestuelles de qualité. L'avènement des dispositifs tels que les Wiimotes ou Kinect offrent à leurs utilisateurs la possibilité d'une plus grande interaction avec les personnages virtuels, et ces derniers trouvent leur place dans de nombreux domaines tels que les jeux vidéos, l'éducation, la simulation ou encore les arts.

Il convient donc d'accroître et d'améliorer les interactions en développant le réalisme et l'autonomie comportementale des humanoïdes virtuels : c'est sur ce point que portera le stage. Plus précisément, nous aborderons l'autonomie sous l'angle de la capacité à maintenir une dynamique des interactions entre plusieurs entités virtuelles, mais également entre une entité virtuelle et un humain. Dans un tel cas, les approches à base de scénarios ou de langages scriptés sont difficilement utilisables car synthétiser des interactions dynamiques implique l'introduction de propriétés telles que l'adaptativité, l'anticipation et la prise de décision en temps réel.

La capacité de prise de décision suivant les intentions et l'environnement est intrinsèque à l'humain. Il est toutefois difficile, dans un monde virtuel, de gérer la complexité d'une interaction humaine et sa modélistion peut être l'objet de différentes propositions. Des approches visent à pourvoir les humanoïdes virtuels d'une autonomie comportementale poussée, à l'aide d' architectures en couches gérant chacune un niveau différent (perception–décision–action). On peut citer par exemple des architectures pour la coordination de mouvement (Multon et coll., 2001, Turvey et coll., 1991) où les PaT-Nets (Badler et Webber, 1995).

D'une manière générale, en simulation de comportement, on trouve trois niveaux : la *cognition*, représentée par les architectures cognitives telles que Soar, ACT-R ; la *dynamique de l'interaction* avec par exemple les architectures pour la coordination de mouvement citées précédemment ; la *gestuelle* avec des techniques comme la cinématique ou la capture de mouvement. La figure 1.1 illustre ces trois niveaux.

La *dynamique de l'interaction* constitue le cadre dans lequel le présent travail se situe. Dans ce contexte, les architectures cognitives sont peu interactives car elles sont basées sur du raisonnement et sont assez lentes (donc peu adaptées à la réactivité). Toutefois, dans le domaine des agents conversationnels animés, SAIBA fait figure d'exception car de récents travaux visent à rendre possible la génération de comportements réactifs dans cette architecture (section 1.2.1). Les techniques pour la gestuelle concernent quant à elles le réalisme des gestes des agents virtuels.

Le domaine de la **psychologie cognitive** a mis en avant un certain nombre d'approches pour la modélisation du comportement humain, que l'on peut séparer en deux familles principales :

1. Les approches basées modèles ;
2. Les approches non-représentationnelles.

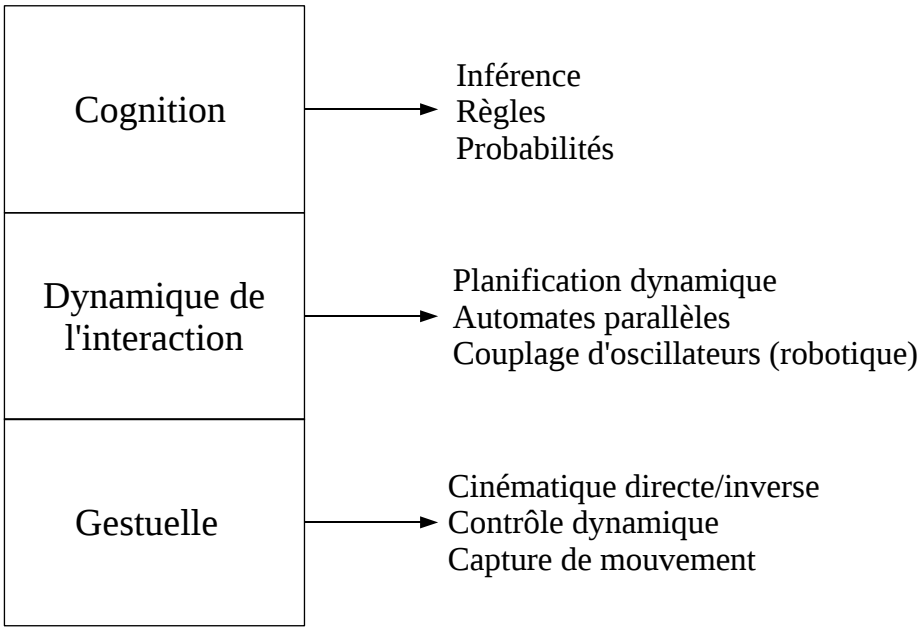


Figure 1.1 – Différents niveaux de modélisation en simulation de comportement et représentation de ces niveaux par des techniques existantes.

Approches basées modèles Il existe des approches dans lesquelles les mouvements sont modélisés comme des plans moteurs décrivant une séquence de commandes musculaires (Keele, 1968). Cependant, les aspects environnementaux n'y sont que très peu pris en compte : au mieux, la notion de perception y est utilisée pour traiter une entrée telle que la position d'une cible à atteindre. Dans l'ensemble, il n'y a pas d'approche agent-environnement dans ce type de modèle.

Des approches orientées « contrôle optimal » existent également (Todorov et Jordan, 2002). Le but de ces approches est de minimiser une fonction objectif ou une fonction de coût satisfaisant des contraintes données. Les trajectoires des mouvements sont ainsi issues de la fonction objectif. Les travaux dans ce type d'approche se focalisent principalement sur la nature de la fonction objectif, comme par exemple la minimisation des tressaillements.

Approches non-représentationnelles Les approches non-représentationnelles mettent quant à elles en avant la dynamique pour décrire le monde, le corps et le couplage sensori-moteur impliqué dans la coordination. Plusieurs approches existent ; parmi elles on peut citer l'approche « perception-action écologique » proposée par Gibson (1986) : l'information perçue dicte le comportement, et il faut alors déterminer la quantité d'informations nécessaires à la génération de comportements aux allures naturelles.

Dans le domaine de l'intelligence artificielle, l'approche par vision active de Ballard (1991) propose par exemple d'exploiter les caractéristiques des images perçues, comme les limites d'une route dans le cas du contrôle d'un véhicule.

La robotique basée sur le comportement initiée par Brooks (1991) propose quant à elle de doter un agent de routines comportementales élémentaires, mais ces systèmes sont purement réactifs et disposent d'un panel de comportements limité.

Des chercheurs proposent désormais des **modèles dynamiques** (Warren, 2006). Afin de comprendre comment les motifs comportementaux sont générés de façon harmonisée avec l'environnement, deux points sont pris en compte :

- la coordination de l'action ;

– la perception de l'environnement.

Les perceptions et les actions sont modélisées par des systèmes dynamiques en intéraction ; **les comportements sont alors modulables et émergent de l'interaction entre l'agent et son environnement**.

Le but de ce stage est d'évaluer la possibilité d'utiliser ces approches, et en particulier celle de Warren (2006) afin d'améliorer les interactions avec des humanoïdes virtuels en donnant à ceux-ci une meilleure adaptabilité à des perturbations pouvant avoir lieu dans leur environnement.

1.2 Agents interactifs temps réel

Cette partie est consacrée à l'introduction de méthodes ayant pour but de renforcer les capacités d'interaction des humanoïdes virtuels. Dans un premier temps, nous présentons quelques techniques utilisées en environnement virtuel pour l'interaction d'agents, puis dans un dernier temps, on quitte le monde virtuel pour la robotique en étudiant l'émergence de synchronicités entre deux robots grâce à deux oscillateurs couplés.

1.2.1 Interaction et adaptation chez les humanoïdes virtuels

La réalisation de modèles permettant à un humanoïde virtuel de bénéficier d'autonomie et de capacités d'interaction prend en compte les concepts de *perception*, de *décision* et d'*action*. Les modèles que nous allons présenter ici sont des architectures construites sur plusieurs niveaux. Chacune d'entre elles reprend les concepts ci-dessus, même s'ils ne sont pas nécessairement aussi explicites. La figure 1.2 illustre le cycle de perception–décision–action.

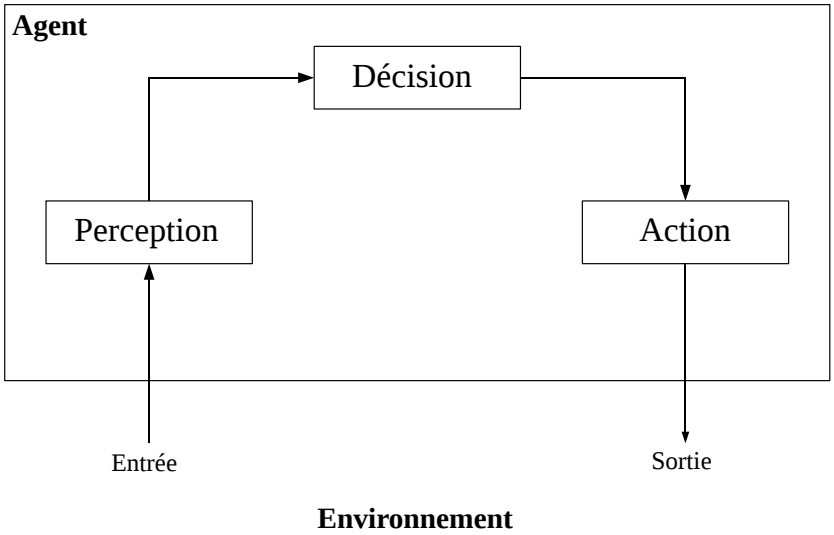


Figure 1.2 – Cycle de perception–décision–action. Les informations perçues permettent à l'agent d'effectuer une prise de décision (en accord avec son but si l'agent est intentionnel) afin de réaliser une action.

Perception : la perception est la prise d'information depuis l'environnement. Ces informations sont de nature sensorielle, comme des informations visuelles, haptiques...

Décision : suivant les informations perçues, l'agent doit décider de l'action à entreprendre. Les agents intentionnels ne prendront pas leurs décisions seulement en fonction de leur perception de l'environnement, mais aussi en fonction de leurs intentions (buts à atteindre).

Action : l'agent exécute l'action déterminée. Cette action peut modifier l'environnement.

La représentation du cycle de perception–décision–action n'est qu'une structure générale pour la réalisation de personnages virtuels autonomes et interactifs. Cependant, l'étude de divers travaux a montré l'omniprésence de cette structure générale, bien qu'à un grain plus fin, des différences significatives existent dans chaque architecture.

Études de cas

Coordination de mouvements Multon et coll. (2001) proposent une architecture basée sur la planification et des techniques d'intelligence artificielle pour la coordination de mouvements appliquée à un jongleur pouvant jongler avec un nombre dynamique de balles.

Le comportement du jongleur n'est pas basé sur un but global mais sur une décomposition d'actions complexes en séquences d'actions élémentaires. Ces actions élémentaires sont définies sur un intervalle de temps pouvant évoluer suite à des événements externes. Dans le cas du jongleur, les actions élémentaires sont :
- lancer une balle ;
- attraper une balle.

Les actions élémentaires sont composées de pré/post conditions, modélisées par des prédicats valant *faux* avant l'exécution de l'action et *vrai* après.

Les bras sont des unités motrices concurrentes exécutant les actions élémentaires. Le choix de l'unité motrice requise pour exécuter une action est effectuée par un niveau de coordination : ce dernier détermine, selon un taux d'occupation (sur un intervalle de temps) de l'unité motrice, et un critère de succès (distance au but de l'unité motrice), quelle est l'unité motrice la plus adaptée à produire l'action.

Le modèle se compose donc des trois couches suivantes :

1. Haut-niveau : décomposition d'actions complexes en actions élémentaires et sélection de l'action à effectuer suivant l'état du système ;
2. Niveau de coordination : suivant l'action élémentaire renvoyée par le haut-niveau, sélection de l'unité motrice requise pour exécuter l'action ;
3. Niveau des unités motrices : exécution des actions élémentaires fournies par le niveau de coordination.

Le système a été testé avec un nombre varié de balles. Elles ne peuvent cependant pas être introduites à tout moment dans la simulation sous peine de donner lieu à des étrangetés visuelles ou à des vitesses anormales des bras.

Parallel Transition Networks Badler et Webber (1995) proposent de modéliser le comportement d'humanoïdes virtuels à l'aide de la planification et de PaT-Nets (*Parallel Transition Networks*). L'architecture est en fait composée de deux niveaux : la boucle SCA (*Sense-Control-Action*) correspond au bas niveau de l'architecture tandis que les Pat-Nets et la planification correspondent au haut niveau.

Sense : perception de l'environnement (objets, autres agents, topologie du terrain) ;

Control : réception des signaux sensoriels puis sélection d'un type de contrôle :
- attraction : création d'un signal en direction du signal reçu ;
- répulsion : création d'un signal dans la direction opposée du signal reçu.

Action : exécution d'une action (marche, position des mains, etc) suivant le signal de contrôle reçu.

Un agent muni seulement d'une boucle SCA sera purement réactif. La structure de contrôle de plus haut niveau obtenue grâce aux PaT-Nets et à la planification permet d'obtenir des agents délibératifs.

Les PaT-Nets sont des automates parallèles surveillant la boucle SCA et adaptent le comportement de l'agent suivant ses buts et sa perception de l'environnement. Ils sont composés d'états représentant une action à réaliser et de transitions permettant le passage d'un état à un autre.

La planification permet de prévoir la prochaine action à réaliser dans la réalisation d'une tâche. La planification dépend surtout de ce que l'on souhaite faire faire aux agents du système.

Un exemple de système à base de planification et de PaT-Nets est le système Gesture Jack (Cassell et coll., 1994). Ce système consiste en deux agents conversationnels : l'un à un but à atteindre tandis que l'autre doit l'aider à y parvenir. Ici le planificateur gère la génération des dialogues selon une base de faits contenant des informations sur le monde, les buts des agents, les croyances des agents (à la fois qu'ils ont du monde et qu'ils ont l'un de l'autre). Les PaT-Nets sont utilisés pour la synchronisation du dialogue et de la gestuelle des agents (mains et regards). Il y a dans le système deux types de PaT-Nets : un pour la gestuelle, l'autre pour le regard. Les mouvements sont alors prévus selon le contexte.

Introduction de comportements réactifs dans l'architecture SAIBA SAIBA (Bevacqua et coll., 2009) est une architecture pour la génération de comportements, modélisée sur trois niveaux (figure 1.3).

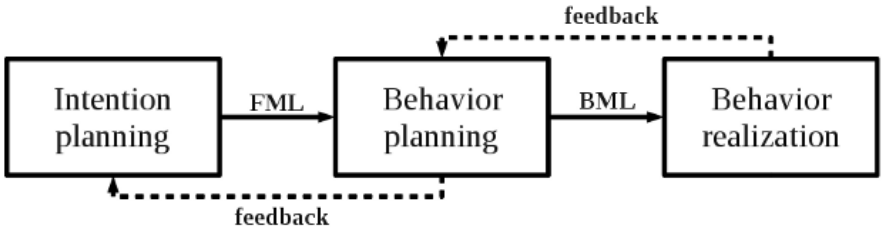


Figure 1.3 – Architecture SAIBA (Bevacqua et coll., 2009)

Le planificateur d'intention détermine les buts courant de l'agent, et les encode en FML (*Function Markup Langage*). Les intentions de l'agent sont ensuite prévues par le planificateur de comportement qui les encode en BML (*Behavior Markup Langage*) et qui seront réalisées par le module de réalisation de comportement.

Estimant que FML seul ne pouvait rendre compte de toutes les interactions pouvant avoir lieu entre humains, les auteurs ont décidé de compléter l'architecture afin de rendre possible des comportements réactifs lors d'une interaction. Ces extensions font que SAIBA se situe maintenant à la limite du niveau cognitif et du niveau de la dynamique de l'interaction (voir figure 1.1).

Les modifications ont principalement lieu dans le module de planification de comportement : celui-ci comprend désormais les trois sous-modules visibles sur la figure 1.4.

Le but de cette extension est de permettre aux agents de réagir à de nouveaux événements (prévus ou non) survenant dans l'environnement. Lorsqu'un tel événement survient, le module de comportement réactif (*Reactive behavior*) va calculer une réaction immédiate. Le module de sélection (*Action selection*) va soit sélectionner l'action calculée par le module de comportement réactif, soit sélectionner l'action planifiée précédente. Le choix s'effectue grâce à des priorités sur chaque action, déterminées suivant les intentions de l'agent. Une replanification dynamique du comportement est possible par le module de représentation FML (*FML chunked representation*).

L'utilisation de morceaux de fichiers FML (*FML chunk*) doit permettre au système de générer des comportements en temps réel : un fichier FML trop important va engendrer des temps de

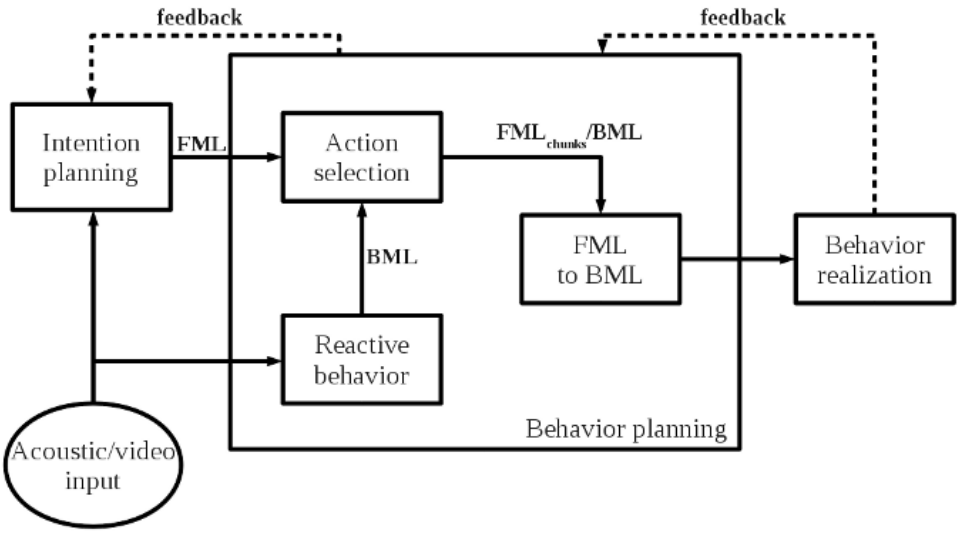


Figure 1.4 – Extension de l'architecture SAIBA. Le module de planification de comportement est composé de plusieurs sous-modules permettant de générer des comportements réactifs.

réponse plus long de la part de l'agent, ce qui est en contradiction avec la notion de comportement réactif.

Les trois modèles présentés dans cette partie ont montré l'introduction de dynamique de l'interaction dans des agents virtuels grâce à différentes méthodes. La question est de savoir s'il est possible pour des comportements d'émerger d'un système sans qu'il soit nécessaire au préalable de les avoir explicités (à l'aide de prédicats, de modules, etc). Ce point est l'objet du prochain paragraphe, traitant de l'émergence de synchronicités chez des robots.

1.2.2 Émergence de synchronicités chez des robots

Le *turn-taking* est caractéristique de la communication entre humains. Ce phénomène consiste en un échange (de parole par exemple) entre deux personnes (ou plus) alternativement. Prepin et Revel (2007) exploitent cette caractéristique dans le domaine de la robotique. Leur modèle est basé sur deux oscillateurs couplés desquels va émerger une synchronisation entre les agents : la provoquation de synchronisation chez les robots revient à faire en sorte d'une part que chaque robot soit sensible au comportement de l'autre et d'autre part qu'il puisse influencer l'autre robot par son propre comportement.

Bien qu'appartenant au domaine de la robotique, ce modèle et ses résultats se rapprochent le plus de la problématique soulevée dans le cadre de ce travail, à savoir l'émergence de comportement adaptatif à l'aide de systèmes dynamiques.

Le système est constitué de deux robots. Chacun d'entre eux est équipé d'un bras avec une articulation (une épaule). Un bras est piloté par un oscillateur constitué de trois neurones ayant la capacité de s'inhiber. Lorsque l'oscillateur est « haut », il provoque un mouvement du bras. S'il est « bas », le bras s'immobilise.

Si une perturbation survient dans l'oscillateur, celui-ci va conserver la phase induite par la perturbation et sa fréquence revient à la fréquence d'origine à l'arrêt de la perturbation (figure 1.5).

Les robots sont également munis d'une caméra, chacun voyant l'autre. De cette manière, si l'un voit que l'autre est en mouvement, il va inhiber sont propre oscillateur : il se produira ainsi des alternances de mouvements et de périodes d'attente.

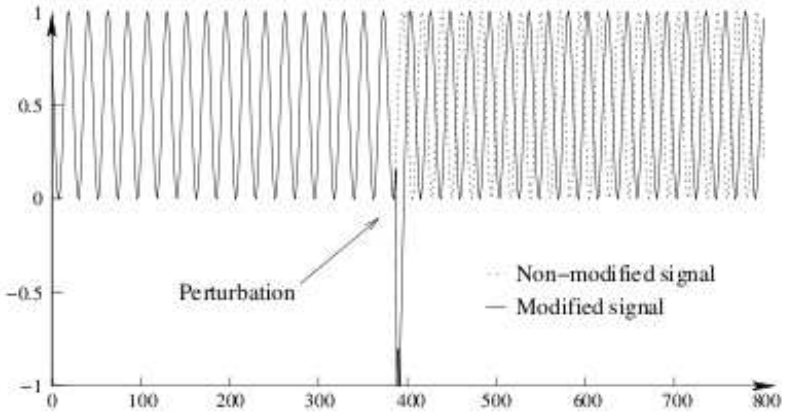


Figure 1.5 – Perturbation d'un oscillateur : la phase induite par la perturbation est conservée et la fréquence d'origine est retrouvée à l'arrêt de la perturbation.

La caméra est donc le module de perception du robot : si un mouvement est perçu, un signal d'inhibition est envoyé à l'oscillateur et l'activité du bras cesse. L'activation motrice est quant à elle résultante de l'activation de l'oscillateur.

La figure 1.6 illustre l'émergence du *turn-taking* au sein du système : après une phase d'amorce, les deux robots se stabilisent dans une anti-synchronisation.

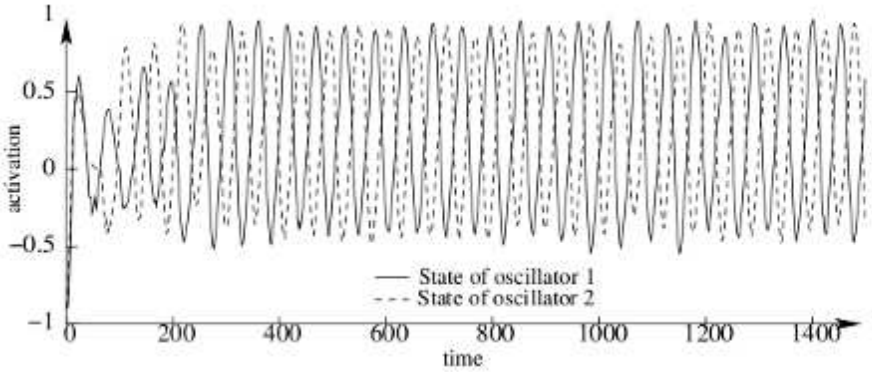


Figure 1.6 – Stabilisation du système : lorsque les oscillateurs s'activent, les robots commencent par bouger ensemble puis les oscillateurs s'anti-synchronisent après un certain temps.

Cet exemple montre un usage de système dynamique appliqué à l'interaction entre deux agents. Chaque agent a la capacité d'influencer l'autre et les oscillateurs sont sensibles aux perturbations pouvant survenir dans l'environnement.

La section suivante porte sur la modélisation comportementale dans les sciences cognitives, et en particulier sur le modèle proposé par Warren (2006), basé sur les sytèmes dynamiques.

1.3 Dynamique du comportement en psychologie cognitive

Le domaine de la psychologie cognitive a mis en avant un certain nombre d'approches pour modéliser les comportements, comme vu dans la section 1.1. Le modèle proposé par Warren

(2006) est basé sur les systèmes dynamiques et met en avant la notion de perception et d'action dans le cadre des interactions agent–environnement. Après une introduction aux systèmes dynamiques, le modèle de Warren est présenté. Nous verrons également comment l'émergence de comportement a lieu avec un tel modèle, et quelles sont les limites de cette approche.

1.3.1 Introduction aux systèmes dynamiques

Les systèmes dynamiques permettent de décrire l'évolution d'un système au cours du temps. Les quelques définitions suivantes introduisent des notions nécessaires à la compréhension de ces systèmes :

Définition 1 *Les* variables d'état *sont un ensemble de variables décrivant un système dynamique.*

Définition 2 *Une* loi dynamique *décrit l'évolution des variables d'état (et donc du système) au cours du temps. La loi dynamique d'un système est donné par un ensemble d'équations différentielles.*

Définition 3 *L'*espace d'états *est l'ensemble de toutes les valeurs possibles des variables d'état. Le comportement du système correspond aux changements des variables d'états et il peut être représenté par des trajectoires dans l'espace d'états.*

Définition 4 *Attracteurs, répulseurs et bifurcation :*
- *Un* attracteur *est une location dans l'espace d'états vers laquelle les trajectoires* convergent*. Si l'état d'un système est perturbé, pour peu que la distance de l'attracteur soit suffisamment petite, la loi dynamique permet de ramener l'état à l'attracteur.*
- *Un* répulseur *est une location dans l'espace d'états de laquelle les trajectoires* divergent*. En cas de perturbation dans le système, même si celle-ci n'a éloigné l'état que d'une distance infinitésimale du répulseur, la loi dynamique l'en éloignera.*
- *Une* bifurcation *est un changement dans le nombre ou la stabilité des attracteurs et des répulseurs.*

Définition 5 *Un* paramètre *est un terme changeant sur une échelle de temps plus longue que les variables d'état.*

1.3.2 Dynamique comportementale (Warren, 2006)

L'approche proposée par Warren (2006), la *dynamique comportementale*, est une approche dirigée vers le but et est basée sur l'utilisation des systèmes dynamiques avec le postulat suivant : dans le système dynamique, les solutions comportementales stables (autrement dit le but à atteindre) sont les attracteurs et les transitions entre comportements sont des bifurcations. Le système global est composé de deux systèmes dynamiques couplés : *l'agent* et *l'environnement.*

Environnement L'équation suivante caractérise le système dynamique correspondant à l'environnement :

$$\dot{e} = \Phi(e, f)$$

Où e est un vecteur de variables d'état de l'environnement, f un vecteur de forces extérieures. Φ sont les lois de la physique gouvernant l'environnement. Un changement dans l'environnement correspond à une fonction de son état courant en fonction des forces extérieures agissant sur lui.

Agent L'agent est quant à lui caractérisé par l'équation :

$$\dot{a} = \Psi(a, i)$$

Où a est un vecteur de variables d'états de l'agent qui décrit l'état courant du *système d'action* et i est un vecteur de variables informationnelles. La loi dynamique (ou loi de contrôle) de l'agent est donnée par Ψ. On définit une action par une fonction de l'état courant du système d'action avec des variables informationnelles (c'est-à-dire l'action à effectuer suivant les informations perçues de l'environnement).

Couplage de l'agent et de l'environnement Deux fonctions sont utilisées pour coupler les deux systèmes dynamiques correspondant à l'agent et à l'environnement. La première, la fonction effecteur, transforme le vecteur de variables d'action en activations musculaires produisant des forces dans l'environnement. La fonction effecteur est définie comme suit :

$$f = \beta(a)$$

La seconde fonction utilisée pour le couplage est la fonction d'information qui transforme les propriétés du système agent-environnement en des variables informationnelles :

$$i = \lambda(e)$$

Le système dans sa globalité (c'est-à-dire les deux systèmes agent et environnement ainsi que leur couplage) est représenté sur la figure 1.7.

La dynamique de la perception et de l'action se voit comme un cycle : l'agent perçoit des informations de l'environnement. Ces informations peuvent être optiques, acoustiques, haptiques... L'agent va ensuite effectuer une action et les changements dans les variables du système d'action sont susceptibles d'induire des changements dans les variables informationnelles de l'environnement ; il faut donc mettre à jour l'état courant du système agent-environnement.

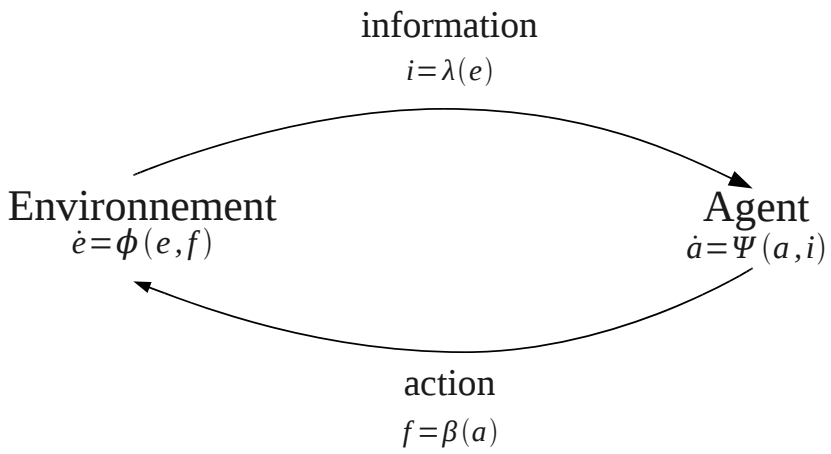


Figure 1.7 – Dynamique de la perception et de l'action

La dynamique comportementale du système est alors définie, à un plus haut niveau, par l'équation suivante :

$$\dot{x} = \Omega(x, s)$$

Où x est un vecteur de variables d'état du comportement et s un vecteur de paramètres du système. De cette manière, les buts correspondent à des attracteurs dans l'espace d'état. Du fait des interactions entre l'agent et l'environnement, les emplacements des attracteurs et des répulseurs dans l'espace d'état peuvent être sujet à des modifications. Les bifurcations de l'agents peuvent alors permettre de nouveaux comportements.

Émergence de comportements

L'usage de systèmes dynamiques permet l'émergence de comportements résultants du couplage entre l'agent et l'environnement. L'interaction de ces deux composants crée un champ de vecteurs dont les attracteurs correspondent à des solutions stables de la tâche. L'« exploration » du champ de vecteurs (autrement dit la recherche d'une solution) par l'agent lui permet de percevoir le résultat de ces actions, conduisant ainsi à un retour permettant d'affiner la recherche d'un comportement stable.

Warren évoque trois niveaux auxquels la dynamicité d'une tâche doit être envisagée :
- La *dynamique du graphe* correspondant aux relations fonctionnelles dans les composants du système et déterminant un régime dynamique (autrement dit la fonction définissant le système dynamique).
- La *dynamique des paramètres* correspondant aux changements dans les paramètres de cette fonction : les attracteurs peuvent alors changer de place dans l'espace ou encore faire faire une bifurcation au système.
- La *dynamique d'état* correspondant à l'évolution du système depuis ces conditions initiales jusqu'à un attracteur.

L'apprentissage d'un comportement par un agent passe par l'obtention d'un régime dynamique qui dépend du but à atteindre. Par exemple, lors d'une tâche consistant à faire rebondir une balle sur une raquette, le régime dynamique correspond à un régime oscillatoire qui est atteint par l'agent grâce aux contraintes physiques de la tâche.

La modification des paramètres (tel que la hauteur du rebond) du système permet de faire évoluer ce dernier afin d'en découvrir les états stables et instables. L'agent procède donc à l'exploration d'espaces de variables de contrôle et d'information lui permettant ainsi de déterminer la loi dynamique de la tâche :
1. L'obtention d'un régime dynamique limite la dimensionnalité de l'espace des variables de contrôle, permettant ainsi à l'agent d'identifier rapidement les variables utiles (c'est-à-dire les variables dont les changements sont immédiatement répercutés sur la stabilité de la tâche). Un exemple de variable de contrôle est la phase oscillatoire de la raquette.
2. La dimension de l'espace des variables informationelles est limitée par les contraintes imposées par la tâche. Une telle variable est par exemple la trajectoire de la balle dans le cas du rebond de balle.

L'émergence de comportement est donc réalisée par l'exploration des espaces de variables de contrôle et d'information, dont les dimensions sont réduites du fait de la spécificité de la tâche à accomplir. De plus, la perception de la stabilité ou de l'instabilité dans le champ de vecteur permet l'établissement de la loi dynamique de la tâche.

Modéliser le comportement avec cette approche requiert d'identifier avec soin les variables utiles pour l'exécution d'une tâche, de déduire les équations du système, mais rencontre aussi des limites en termes de but et d'informations disponibles : Warren a étudié ces limitations et émis des hypothèses pour chacune d'elles.

1.3.3 Limitations de la dynamique comportementale

Séquences d'actions La réalisation d'une tâche passe souvent par la réalisation de séquences d'actions (sous-tâches). Warren illustre bien ce problème avec une tâche pourtant très simple : faire un sandwich au beurre de cacahuète : il faut rassembler les ingrédients, disposer une tranche de pain, ouvrir le pot, etc. De plus, ces actions peuvent être liées entre elles par des dépendances (ouvrir le pot avant de pouvoir se servir).

De tels comportements seraient hiérarchisés par une représentation interne pré-existante de la séquence d'actions. Mais la variabilité observée par les psychologues peut également laisser penser que les séquences sont réalisées dynamiquement lors de la tâche.

Une approche proposée par Keijzer (1998) est l'étude des changements des régimes dynamiques au cours du temps. Pour cela, l'utilisation de systèmes dynamiques multi-échelles couplés

permettent de modéliser des interactions entre agent et environnement se déroulant sur plusieurs échelles de temps. Couplés à des réseaux de neurones, les systèmes sur une courte échelle de temps sont ensuite couplés à des systèmes dont l'échelle est plus grande. Le comportement global modifie les réseaux de neurones afin de permettre un apprentissage.

Toutefois, les comportements séquentiels restent encore à étudier et une piste intéressante pour poursuivre dans cette voie serait l'utilisation de réseaux de neurones récurrents.

Comportements anticipatifs Warren définit les comportements anticipatifs comme étant des actions ne dépendant pas seulement des informations actuelles, mais également d'un but « distant ». Aussi, on retrouve dans ce cas de figure la notion de multi-échelles proposée par Keijzer. Il faut parvenir à faire émerger une trajectoire à long terme dans l'espace d'état à partir des interactions agent/environnement à court terme.

Les paramètres de contrôle de la dynamique à court terme permettent d'aller vers des trajectoires à plus long terme, et à l'inverse, la dynamique à long terme permet de modifier les interactions à court terme grâce à la disponibilité de nouvelles informations par exemple.

Comportements prédictifs Le comportement prédictif est lui définit par des actions dépendant de propriétés cachées de l'environnement, telles que la masse d'un objet, sa fragilité, etc. Dans ce cas, l'idée est que l'agent doit avoir une représentation interne des objets comprenant également leurs propriétés. L'agent utiliserait donc ces représentations internes avec d'autres informations disponibles (visuelles par exemple) afin de prédire l'évolution des entités avec lesquelles il interagit.

Une modification des propriétés d'un objet impliquerait une adaptation de l'agent. L'apprentissage des liens entre un objet et ses propriétés cachées doit se faire par l'expérience passée de l'agent de façon à pouvoir les incorporer dans une loi dynamique par exemple.

De nouveau, il ressort donc des comportements prédictif la notion d'interactions entre agent et environnement sur plusieurs échelles.

Comportements stratégiques Le dernier type de comportement étudié par Warren est le comportement stratégique : lors de ses interactions avec l'environnement, l'agent peut être amené à prendre en compte l'historique de cet environnement ou encore des statistiques sur les événements. L'apprentissage de ces faits est possible lors des interactions entre l'agent et l'environnement et, encore une fois, a une influence sur la dynamique à plusieurs échelles.

1.4 Synthèse

Depuis plusieurs années, des modèles ont été élaborés afin de répondre aux problématiques de l'interaction d'agents virtuels autonomes. Bâties sur la boucle de perception–décision–action, ces modèles explorent différentes voies pour développer la crédibilité des interactions, comme l'utilisation de précicats (Multon et coll., 2001) ou les automates parallèles (Badler et Webber, 1995). Ces modèles utilisent des couches, des symboles, de la planification ; cependant ils ne se focalisent pas en priorité sur la dynamique émergeant de l'interaction.

En revanche, cette émergence apparaît dans le domaine de la robotique : l'approche utilisée par Prepin et Revel (2007) a montré que le couplage de systèmes dynamiques (ici deux oscillateurs) permet l'émergence d'une interaction entre les deux agents du système. Ce système n'a pas nécessité l'utilisation de couches pour gérer les comportements des agents, et n'est pas non plus basé sur la planification d'actions : la dynamique de l'interaction émerge sans que les comportements aient été explicités, même dynamiquement.

En psychologie cognitive, une hypothèse concernant les interactions agent/environnement est que l'humain est adaptatif et utilise les propriétés de stabilité de la tâche dans sa recherche d'efficacité. La réalisation d'une tâche spécifique nécessite l'identification des variables essentielles du système pour atteindre une solution stable. Une telle tâche doit donc avoir fait l'objet d'études en psychologie cognitive pour que nous puissions la modéliser dans un environnement virtuel.

Nous avons choisi de porter notre modélisation sur la tâche consistant à faire rebondir une balle sur une raquette de manière rhytmique. Bien que cette tâche n'implique la présence que d'un agent dans l'environnement, elle a fait l'objet de nombreuses études et est donc adaptée à une modélisation en environnement virtuel permettant l'évaluation de l'approche par couplage de systèmes dynamique dans l'optique de renforcer le réalisme des interactions entre agents.

Partie 2

Cas d'étude : le rebond de balle

Cette partie donne la description des systèmes en jeu dans le cas du rebond de balle ainsi que les clefs pour réaliser le couplage entres ces systèmes.

2.1 Généralités

Faire rebondir une balle sur une raquette dans une dimension (verticale) est un cas d'étude simple pour appréhender la dynamique comportementale et est basé sur le cycle de perception-action suivant : la force appliquée par la raquette sur la balle influe sur l'état de l'environnement (la balle en particulier). Cela génère des perturbations donnant lieu à de nouvelles informations sur la trajectoire de la balle. Ainsi, une modification de la force appliquée par la raquette permet de modifier la dynamique de la balle, et en retour le cycle de la raquette. On retrouve bien ici le cycle de perception/action explicité par Warren (2006) et illustré sur la figure 1.7 page 10.

Les études de psychologie cognitive écologique ont pour objet de comprendre comment un acteur réalisant cette tâche exploite les contraintes de l'environnement afin de réaliser ladite tâche ; ces contraintes peuvent être physiques ou encore informationnelles et doivent permettre à l'acteur de se stabiliser dans un rebond stable, c'est-à-dire de frapper rhytmiquement la balle de façon à ce que la hauteur de celle-ci soit stable.

Il existe deux types majeurs de contrôle : l'un est *passif*, autrement dit le système peut se trouver dans un état stable sans qu'il soit nécessaire de percevoir les informations de l'environnement. L'autre est *actif* : dans ce cas, l'analyse cognitive des informations perçues permet de mener le système dans un état stable.

Rebond passivement stable Schaal et coll. (1996) ont montré que le rebond de balle peut être *passivement stable*, c'est-à-dire que le rebond a lieu indéfiniment et peut compenser de petites perturbations pouvant parvenir dans le système sans contrôle actif de la raquette. Ces perturbations peuvent être de différentes natures, comme la modification de l'accélération de la pesanteur, obtenue grâce à un système de réalité virtuelle.

L'analyse du système de rebond de balle a permis de déterminer sous quelle contrainte physique on peut qualifier le rebond de passivement stable : l'impact doit avoir lieu dans le dernier quart du cycle de la raquette (figure 2.1 p. 16), phase ϕ_r pendant laquelle l'accelération de la raquette a_r est négative. En particulier, on a :

$$\frac{-2g(1+\alpha^2)}{(1+\alpha)^2} < a_r < 0 \tag{2.1}$$

Où g est l'accélération de la pesanteur et $\alpha \in [0;1]$ est le coefficient de restitution, caractérisant l'élasticité de l'impact (autrement dit la perte d'énergie au moment de l'impact).

La capacité à maintenir un rebond stable sans contrôle actif dépend très fortement des valeurs initiales des paramètres du système : le coefficient de restitution, la gravité, la période et l'amplitude de la raquette, la position et la vitesse de la balle.

Les notions de rebond *passif* et *actif* ne sont pas exclusives l'une par rapport à l'autre. En effet, selon Sternad et coll. (2001), les acteurs exploitent cette stabilité passive afin de réaliser la tâche, mais ils stabilisent également *activement* le rebond grâce aux informations perçues depuis l'environnement (Morice et coll., 2007).

Contrôle perceptuel Les informations perçues depuis l'environnement peuvent être de nature *visuelle* (trajectoire de la balle), *haptique* (moment et force de l'impact courant) et *acoustique* (moment et force de l'impact également). Les *variables essentielles* sont exploitées par les acteurs dans leur recherche de stabilité. La question est de savoir, parmi les informations citées, quelles sont celles considérées comme essentielles car permettant de stabiliser le rebond.

Cycles de la balle et de la raquette La figure 2.1 p.16 illustre les cycles respectifs de la balle et de la raquette.

Le cycle de la balle est caractérisée par sa période, mesurée comme étant le temps entre deux impacts successifs. Les informations concernant la balle sont sa vitesse initiale (mesurée immédiatement après l'impact), la hauteur maximale du rebond h_b. L'erreur au rebond ϵ est calculée comme étant la différence entre h_b et la hauteur cible à atteindre h_t. Les variables t_{up} et t_{down} représentent respectivement la demi-période ascendante et la demi-période descendante de la balle.

La trajectoire de la raquette est représenté par des oscillations harmoniques. Son cycle est caractérisé par sa période P_r, mesurée comme étant le temps s'écoulant entre deux positions de hauteur maximale successives. L'amplitude correspond à la distance séparant les positions maximale et minimale de la raquette. Le cycle de la raquette peut être découpé en quatre quartiers correspondant à des phases d'accélération et de décélération ; ϕ_r est la phase de la raquette au moment de l'impact.

2.2 Modes de contrôle

Si de nombreuses études ont portés sur le rebond de balle, et notamment sur les critères permettant de déterminer comment est définie la stabilité passive, peu d'études proposent des hypothèses concernant le contrôle de la raquette entre deux impacts.

Bien que la tâche du rebond de balle soit réalisable en régime passivement stable, il est également possible d'inclure dans un modèle un contrôle perceptuel. Le système, déjà perturbable en régime passivement stable, doit permettre d'offrir une robustesse accrue face aux perturbations pouvant survenir dans l'environnement.

Siegler et coll. (2010) ont avancé l'existence de quatres modes de contrôles. Le premier est le contrôle *purement passif*. Dans ce cas, aucun contrôle n'est exercé sur la raquette pour compenser sa trajectoire ; seules de petites perturbations peuvent être compensées.

Le contrôle *purement actif* est un mode dans lequel un contrôle est exercé à chaque cycle mais dans lequel la propriété de stabilité passive du système n'est pas prise en compte. On peut citer l'algorithme miroir de Buehler et coll. (1994) qui calque la vitesse de la raquette sur celle de la balle : les impacts ont dans ce cas lieu avec une accélération positive.

Le contrôle *hybride* prend en compte la propriété de stabilité passive de la tâche : les petites perturbations sont compensées grâce à cette propriété et les perturbations plus grandes sont compensées activement. De Rugy et coll. (2003) ont utilisé ce mode de contrôle dans leur modèle : le rebond devant avoir lieu à une hauteur constante, passé un certain seuil d'erreur, la période de la raquette est ré-ajustée suivant la période de la balle afin de compenser les perturbations de l'environnement (dans ce cas, des modifications du coefficient de restitution).

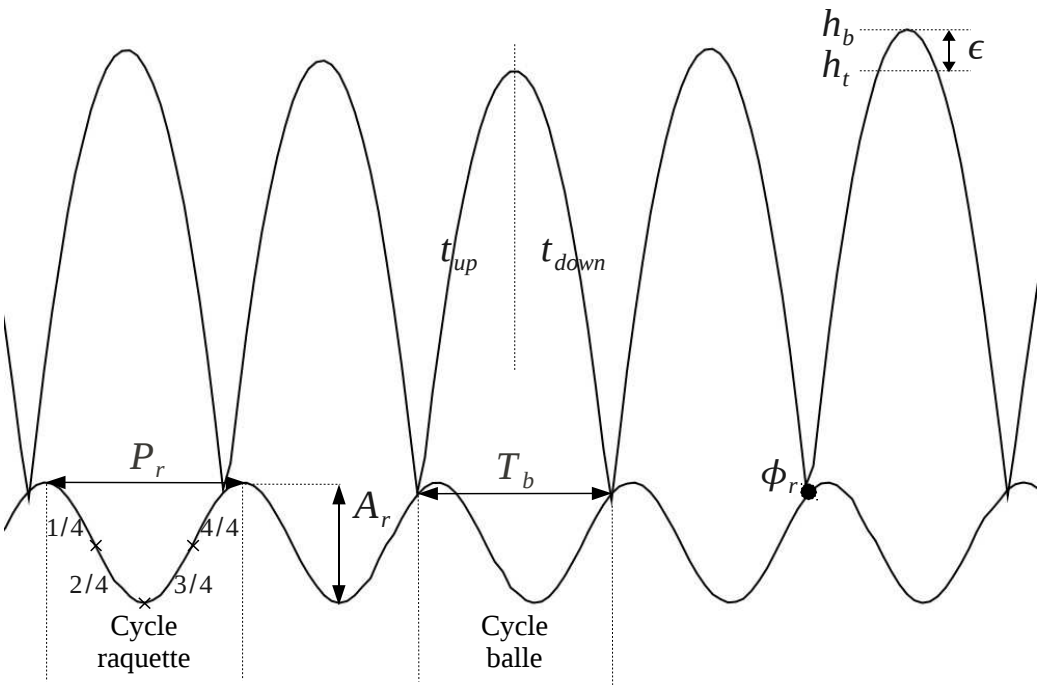


Figure 2.1 – Cycle balle/raquette.

Enfin, le dernier mode de contrôle est le contrôle *mixte*. Dans ce cas également, la propriété de stabilité passive de la tâche est prise en compte. Un contrôle perceptuel est exercé à chaque cycle, ayant pour effet non seulement de compenser les perturbations de l'environnement, mais aussi de maintenir/ramener le système dans son état passivement stable.

Selon Siegler et coll. (2010), le contrôle mixte est utilisé par les acteurs lors de la réalisation de la tâche. Ce mode de contrôle doit permettre de réduire les ajustements ainsi que le temps de relaxation du système, menant ainsi à une plus grande stabilité par rapport aux autres modes de contrôles.

2.2.1 Contrôle actif de la raquette

Dans le cadre de l'approche de la *dynamique de la perception et de l'action* mise en avant par Warren (2006), la mise en place d'un contrôle actif dans un modèle informatique revient à réaliser un **couplage** entre les systèmes représentant respectivement l'agent et son environnement. Le couplage de ces deux systèmes est réalisé par le couplage les variables propres à chaque système, basé sur les observations réalisées en psychologie cognitive écologique sur une tâche donnée – en l'occurence faire rebondir une balle sur une raquette.

Le couplage réalisé donne lieu à un contrôle actif mixte, duquel on doit voir émerger la propriété de stabilité passive de la tâche réalisée, c'est-à-dire une accélération à l'impact négative, dont la valeur se situe dans l'intervalle présenté à la relation (2.1).

Un certain nombre de variables sont suceptibles d'être extraites de l'environnement. Il convient de distinguer les variables essentielles de celles qui ne le sont pas, c'est à dire les variables effectivement utilisées par un acteur réalisant cette tâche.

Informations disponibles depuis l'environnement L'étude de Siegler et coll. (2010) a pour objet de déterminer quelles sont les informations perceptuelles utilisées par les acteurs

réalisant cette tâche. Trois types d'informations sont susceptibles d'être utilisées : *visuelles*, *acoustiques* et *haptiques* (voir paragraphe 2.1).

Sternad et coll. (2001) ont montré que l'utilisation d'informations haptiques et acoustiques seules permettent de réaliser un rebond stable, les informations visuelles aidant à la conservation de la stabilité. Cependant, Siegler et coll. (2010) estiment que les seules informations haptiques et acoustiques ne permettent pas d'apporter les corrections nécessaires suite à une perturbation de l'environnement. En effet, d'après leurs expériences, les perturbations sont compensées en un court laps de temps (la compensation peut avoir lieu en un cycle seulement suivant l'information utilisée).

Informations perçues pour le contrôle actif Les informations visuelles extraites de la trajectoire de la balle semblent être les informations les plus efficaces pour maintenir le rebond en régime stable et apporter les corrections nécessaires en cas de perturbation de l'environnement.

Les résultats de Siegler et coll. (2010) tendent à confirmer la présence d'un contrôle mixte durant la réalisation d'un rebond de balle par un acteur humain. Les informations perçues de l'environnement permettant de mener à bien cette tâche seraient les suivantes : la demi-période ascendante de la balle t_{up}, la demi-période descendante de la balle t_{down} (la période de la balle peut s'exprimer de différentes manières : $T_b = 2t_{up} = 2t_{down}$) et la hauteur maximale du rebond h_b qui permet d'estimer l'erreur au rebond.

Exploitation des informations perçues Les informations visuelles extraites de l'environnement doivent servir à piloter la raquette afin d'obtenir un comportement à la fois stable et adaptatif.

Les oscillations de la raquette sont contrôlées par des paramètres modulant de manière *indépendante* leur période et leur amplitude (Siegler et coll., 2010).

Le contrôle de la période de la raquette est lié à la perception des informations visuelles concernant la période de vol de la balle : suite à l'analyse des données recueillies lors de leurs expérimentations, Siegler et coll. (2010) ont observé que la demi-période ascendante de la balle t_{up} est fortement corrélée à la période des oscillations de la raquette. La demi-période descendante de la balle t_{down} est également corrélée à la période des oscillations de la raquette, vraissemblablement pour effectuer un ré-ajustement. Sachant que la période de la raquette doit approximer la période de la balle, l'ajustement de la période de la raquette a lieu sitôt que t_{up} a été perçue par l'acteur réalisant la tâche.

Si les informations visuelles à propos de la hauteur du rebond devraient permettre le contrôle de l'amplitude, la situation est cependant plus problématique en réalité : en effet, Siegler et coll. (2010) ont bel et bien noté des ajustements lors des expériences réalisées, mais aucune des variables présentées ne semble être directement liée à ces ajustements. Ils ont alors émis l'hypothèse que la modulation de l'amplitude était réalisée sur une échelle de temps plus longue, grâce à la moyenne de l'erreur au rebond sur les rebonds précédents.

2.3 Le système balle/raquette

Les variables permettant de caractériser la balle sont sa position x_b et sa vitesse $\dot{x}_b$. la trajectoire de la balle entre deux impacts suit l'équation suivante :

$$x_b(t) = x_b^+ + \dot{x}_b^+ t - 0.5gt^2 \tag{2.2}$$

Où x_b^+ et $\dot{x}_b^+$ sont respectivement la position et la vitesse de la balle immédiatement après l'impact, et g est l'accélération due à la pesanteur.

Les impacts entre la balle et la raquette correspondent à l'équation suivante :

$$(\dot{x}_b^+ - \dot{x}_r) = -\alpha(\dot{x}_b^- - \dot{x}_r) \tag{2.3}$$

Où $\dot{x}_b^-$ et $\dot{x}_b^+$ est la vitesse de la balle immédiatement avant et après l'impact, $\dot{x}_r$ est la vitesse de la raquette.

La raquette est quant à elle caractérisée par sa période P_r, son amplitude A_r et sa phase ϕ_r. Les détails concernant la modélisation du système sont donnés dans la partie 3.2.

Partie 3

Modélisation du système balle/raquette

3.1 Environnement de simulation

L'implémentation du cas d'étude est réalisé en environnement virtuel. Cet environnement a été créé avec la bibliothèque ARéVi [1], développée au CERV, permettant de réaliser des simulations 3D en temps réel. Le comportement physique des objets dans l'environnement est géré par le moteur physique ODE [2] offrant entre autres la possibilité d'exploiter des propriétés physiques de l'environnement, comme la gravité.

Si les oscillations de la raquette sont exclusivement contrôlées grâce à l'oscillateur présenté dans la section 3.2 ci-dessous, les mouvements de la balle sont entièrement gérés par le moteur physique. Les deux systèmes interagissent dans l'environnement. À aucun moment les équations modélisant le vol de la balle et les impacts avec la raquette ne sont utilisées pour calculer la vitesse ou position de la balle : les informations relatives à l'état de la balle sont perçues dans l'environnement et utilisées pour le couplage (section 3.3).

L'environnement de simulation a été paramétré dans un premier temps afin de pouvoir observer un rebond de balle passif. La mise en place d'un rebond passif maintenu est fortement dépendant des conditions initiales à la fois de l'environnement (gravité, coefficient de restitution), de la balle (vitesse et position initiales) et de la raquette (période et amplitude). Pour les conditions environnementales $g = 9.81\,m.s^{-2}$ et $\alpha = 0.48$, les conditions initiales de chaque système (balle et raquette) choisies sont les suivantes :

- Balle : $x_{b,0} = 0.16\,m$ et $v_{b,0} = 4.6\,m.s^{-1}$
- Raquette : $P_r = 0.66\,s$ et $A_r = 0.18\,m$

3.2 Oscillateur et pilotage de la raquette

Dans l'environnement virtuel, la raquette est pilotée par un oscillateur constitué de deux neurones (figure 3.1) s'inhibant l'un l'autre et produisant alors des schémas oscillatoires. Ce modèle a été proposé par Matsuoka (1985, 1987) et l'activité des neurones du réseau est générée par les équations (3.1).

$$\left.\begin{aligned} T_r\dot{x_i} = -x_i - \sum_{j=1}^{n} a_{ij}y_j + s_i - bf_i \\ T_a\dot{f_i} = -f_i + y_i \\ y_i = g(x_i) \end{aligned}\right\} \tag{3.1}$$

1. Atelier de Réalité Virtuelle, http ://svn.cerv.fr/trac/AReVi, sous licence LGPL
2. Open Dynamics Engine, http ://www.ode.org/, sous licence LGPL ou BSD

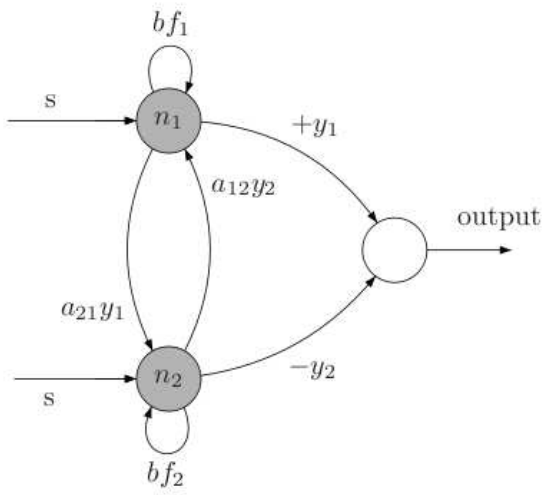


Figure 3.1 – Oscillateur constitué d'un réseau de deux neurones.

Où n est le nombre de neurones dans le réseau, s est une entrée externe au réseau, x_i est l'état interne du neurone i, f_i est le degré d'auto-inhibition du neurone i modulé par b, a_{ij} est le poids connectant le neurone j au neurone i, y_i est la sortie du neurone i, $g(x) = max(0, x)$, et T_r et T_a sont deux constantes de temps.

Seuls deux paramètres dans ce système sont fixés et constants durant toute simulation : a et b, permettant de réguler l'inhibition d'un neurone à l'autre ou d'un neurone sur lui même. Ici ces paramètres ont pour valeur : $a = 1.4$ et $b = 2.5$. Ces valeurs permettent aux neurones de bénéficier d'une inhibition suffisamment grande pour qu'un changement d'amplitude modérée soit répercuté rapidement sur les oscillations : en effet, avec un oscillateur de ce type une modulation de période ou d'amplitude ne se produit pas en un temps égal à zéro (ce qui est idéal ici, puisque même si le focus n'est pas fait sur la dynamique du geste, l'apparence des oscillations doivent avoir une cohérence par rapport à un geste humain). L'adaptation aux nouvelles valeurs de période ou d'amplitude injectées dans l'oscillateur doit toutefois être répercutée suffisamment rapidement à cause du fait que l'on réalise un contrôle actif ayant lieu à chaque cycle : une modulation à une oscillation doit être répercutée sur l'oscillation la suivant immédiatement.

Si on souhaite injecter un changement dont l'amplitude est plus importante, l'adaptation prendra cependant un temps plus grand pour peu que a et b aient toujours les valeurs 1.4 et 2.5. Ces valeurs constituent donc un compromis permettant de ne pas causer d'aberration dans les oscillations lorsque l'on cherche à les moduler, mais qui a ses limites en terme d'amplitude de changement.

Contrôle de la période La période de l'oscillateur est contrôlée avec les deux paramètres T_r et T_a :

$$\begin{aligned} T_r &= P_t * c_1 \\ T_a &= P_t * c_2 \end{aligned} \tag{3.2}$$

P_t est la période cible à laquelle on souhaite fixer l'oscillateur. Les coefficients c_1 et c_2 ont pour valeurs respectives 0.1456 et 0.363 dans l'environnement virtuel. Ces valeurs ont été choisies de manière à ce qu'en sortie, la période effective de l'oscillateur corresponde à la période en entrée.

Contrôle de l'amplitude L'amplitude de l'oscillateur est quant à elle modulée par le paramètre s :

$$s = A_t * \beta \tag{3.3}$$

Où A_t est l'amplitude que l'on souhaite fixer à l'oscillateur. Le coefficient β a pour valeur 1.219 dans l'environnement virtuel. La valeur de β a été choisie afin d'obtenir une amplitude effective de l'oscillateur égale à l'amplitude en entrée.

Cet oscillateur permet donc un contrôle indépendant de la période et de l'amplitude, conformément aux hypothèses avancées par Siegler et coll. (2010).

3.3 Algorithme de contrôle

L'algorithme de contrôle est en fait la réalisation du couplage entre les systèmes de la balle et de la raquette. Ce couplage a été réalisé de manière à ce que les propositions de Siegler et coll. (2010) soit respectées : en particulier, le contrôle réalisé est *mixte*, c'est-à-dire qu'il a lieu à chaque cycle et permet non seulement de corriger les erreurs dues à des perturbations survenant dans l'environnement, mais également de maintenir ou ramener le système dans son état passivement stable.

Il est possible d'utiliser les équations modélisant le vol et l'impact de la balle, présentées au paragraphe 2.3 afin de connaître l'état de la balle au cours du rebond. Dans le cas présent, les équations ne sont pas utilisées, la trajectoire de la balle et les impacts sont entièrement gérés par le moteur physique et l'état de la balle est connu grâce à un module de perception qui permet de connaître les informations nécessaires au contrôle actif, en particulier la période de la balle.

Le couplage réalisé permet de moduler la période et l'amplitude de l'oscillateur pilotant la raquette. Pour cela, les informations nécessaires sont tout d'abord extraites de l'environnement puis utilisées pour moduler période et amplitude de l'oscillateur. Celui-ci pilote la raquette, qui va influencer la trajectoire de la balle et le cycle recommence. Le couplage décrit est illustré sur la figure 3.2 et le détail des ajustements pour la période et l'amplitude de l'oscillateur est décrit dans les paragraphes 3.3.1 et 3.3.2.

La modulation de la période de la raquette seulement ou de l'amplitude de la raquette seulement ne permet pas d'obtenir un rebond stable sur une longue période de temps.

Pour éviter de donner lieu à des comportements de raquette aberrants (vitesse trop élevée, amplitude quasi nulle), l'amplitude A_r de la raquette ne peut être inférieure à 0.08m et sa période P_r ne peut être inférieure à 0.3s. Bien qu'on ne se focalise pas sur la gestuelle, ces valeurs ont été choisies afin d'obtenir une visualisation plus réaliste dans l'environnement.

3.3.1 Modulation de période

La modulation de la période de la raquette est réalisée dès que la demi-période ascendante t_{up} de la balle a été perçue. La période perçue est alors injectée dans l'oscillateur : $P_r = 2t_{up}$.

Un second ajustement est possible en fin de cycle de balle puisqu'à ce moment là, t_{down} et T_b ont été perçus.

Les résultats présentés partie 4 sont réalisés avec d'une part un ajustement unique avec t_{up} et d'autre par avec un second ajustement réalisé à la fin du cycle de la balle lorsque T_b a été perçu.

3.3.2 Modulation d'amplitude

La modulation de l'amplitude doit d'une part permettre de conserver une phase à l'impact cohérente avec les critères de stabilité passive de la tâche et d'autre part permettre de conserver une hauteur de rebond régulière. Cependant, Siegler et coll. (2010) n'ont pas identifié de variable liée à la balle permettant de contrôler l'amplitude : bien que celle-ci doive être modulée de façon à conserver une hauteur de rebond constante, la corrélation avec l'erreur à la hauteur du rebond est très faible. Dans le cas présent, l'hypothèse de Siegler et coll. (2010) selon laquelle l'amplitude

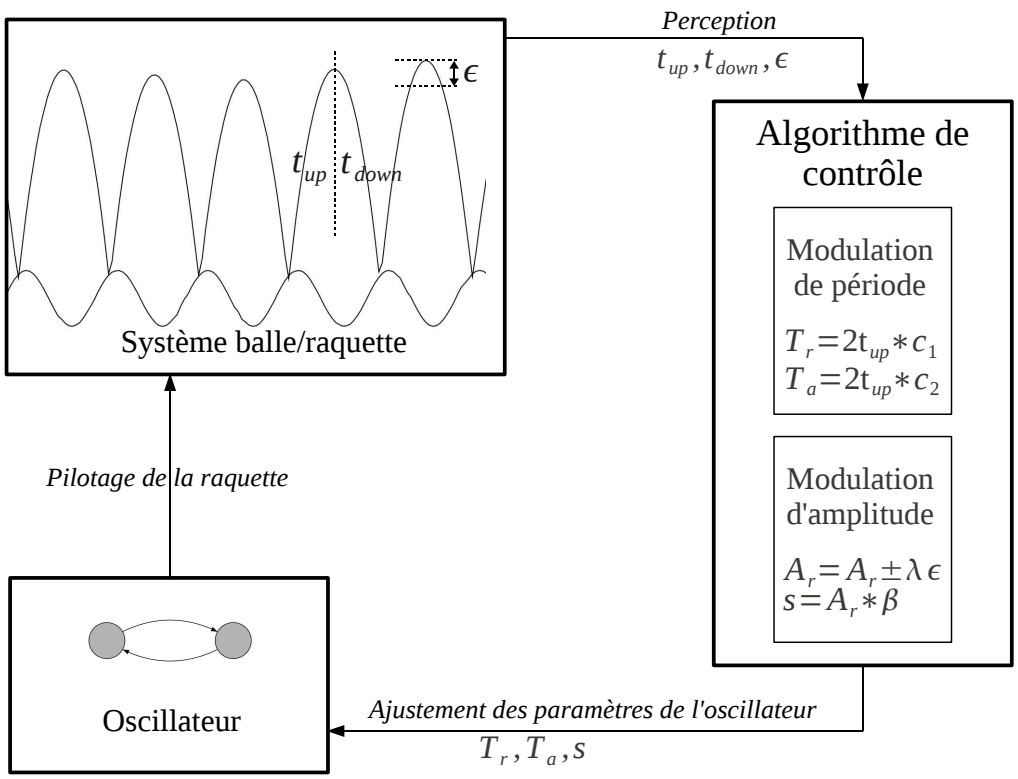


Figure 3.2 – Couplage des systèmes dynamiques avec l'algorithme de contrôle

est modulée grâce à une moyenne de l'erreur au rebond sur les cycles précédents a été retenue. En outre, les questions suivantes n'ont pas trouvé réponse dans l'étude citée, et il a donc été nécessaire d'effectuer des choix :

- Combien de cycle(s) faut-il prendre en compte pour mesurer l'erreur moyenne ?
- Comment moduler l'amplitude en fonction de l'erreur mesurée ?
- À quel moment dans le cycle doit intervenir la modulation de l'amplitude ?

Mesure de l'erreur moyenne Le nombre de cycles retenu pour la mesure de l'erreur moyenne est 4 : autrement dit, à chaque cycle, l'erreur moyenne des 4 derniers rebonds est retenue pour réaliser l'ajustement de l'amplitude de la raquette. Cependant, le fait de travailler avec un moteur physique (et donc un environnement bruité) fait que, même dans le cas du rebond passif :

- La hauteur au rebond ne sera jamais constante au delà d'un certain seuil (on considèrera par exemple une hauteur constante à $\pm 0.2\,m$).
- Des rebonds « aberrants » peuvent se produire, c'est-à-dire des rebonds dont la hauteur est disproportionnée par rapport aux autres rebonds de l'essai.

Une idée est alors de tenir compte de la sensibilité de la moyenne aux valeurs extrèmes : lors du calcul de la moyenne sur les 4 derniers rebonds, on écarte la plus grande et la plus petite valeur. C'est la raison pour laquelle nous proposons de prendre en compte 4 rebonds.

Modulation de l'amplitude suivant l'erreur mesurée Les ajustements de l'amplitude doivent bien entendu être relatifs à l'erreur mesurée. Une manière de moduler les changements d'amplitude de la raquette selon l'erreur retenue est d'utiliser la relation suivante :

$$A_r = A_r \pm \lambda\epsilon \tag{3.4}$$

Où ϵ est l'erreur mesurée et λ un coefficient multiplicateur. Si $\epsilon > 0$ alors le rebond a dépassé la hauteur cible et il faut diminuer l'amplitude. Si $\epsilon < 0$ alors le rebond est trop petit et il faut augmenter l'amplitude.

Le coefficient λ est à déterminer. Les simulations en contrôle actif sont réalisées avec les valeurs suivantes afin de cerner l'influence de ce paramètre : $\lambda = 0.2,\ 0.4,\ 0.6,\ 0.8$. Les ajustements de l'amplitude sont donc plus importants pour une valeur plus grande de λ.

Quand, dans le cycle, ajuster l'amplitude de la raquette ? Le moment pendant lequel doit intervenir la modulation de l'amplitude a été choisi par rapport au comportement de l'oscillateur pilotant la raquette : une modulation à un moment inapproprié durant le cycle peut conduire à des ruptures dans les oscillations. Les ajustements de l'amplitude se produisent alors à la fin du cycle de la raquette.

Partie 4

Résultats & Discussion

Tous les résultats/courbes présentés dans cette partie sont issus de simulations réalisées **dans l'environnement virtuel**. Les perturbations environnementales sont des modifications de l'accélération de la pesanteur g, intervenant lorsque la balle a atteint sa hauteur maximale h_b ; le coefficient de restitution α est constant pour toutes les simulations réalisées et vaut 0.48.

4.1 Rebond passif

L'objectif de ce travail est de mettre en place un couplage entre les systèmes correspondant à la balle et à la raquette grâce aux informations présentées en section 3.3 afin d'observer les propriétés émergentes du système dans sa globalité. Néanmoins le rebond de balle passif exhibe la propriété intéressante du maintien d'un rebond constant, y compris suite à l'introduction de perturbation dans l'environnement virtuel, et ce malgré l'absence de contrôle dans le système.

4.1.1 Méthode

Deux séries de simulations sont réalisées :

1. Les conditions environnementales sont constantes au cours d'une même simulation (la gravité initiale n'est pas sujette à modification).
2. Une perturbation de l'environnement est introduite au cours d'une même simulation (la gravité est modifiée pendant la simulation).

Chaque simulation dure entre 50 et 60 secondes, et pour chaque condition environnementale, une session de cinq simulations est réalisée. Les moyennes et écarts types des variables du système de la raquette sont calculés pour chaque condition sur l'ensemble des cinq simulations de la session.

Les données résultantes des sessions de simulations sont présentées de différentes manières : pour une simulation de la session, les trajectoires de la balle et de la raquette ainsi qu'un portrait de phase du mouvement de la raquette lors de cette simulation (exprimant la vitesse de la raquette en fonction de sa position) permettent d'obtenir un aperçu du déroulement des simulations de la session. Pour l'ensemble de la session, on présente également un tableau de résultats contenant les valeurs moyennes et les écarts types des variables du système de la raquette.

4.1.2 Conditions environnementales constantes

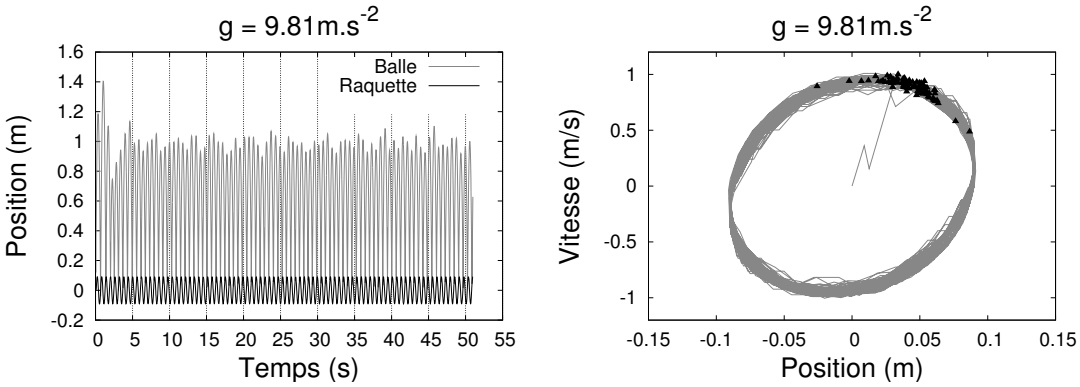
Conditions de références

Dans cette série de simulations, la gravité g vaut $9.81\, m.s^{-2}$. Il s'agit d'une session « témoin », dont les conditions environnementales sont celles ayant servi à calibrer le système, c'est-à-dire

déterminer les valeurs initiales des paramètres des systèmes de la balle et de la raquette pour obtenir un rebond stable sans présence de contrôle actif (section 3.1).

La figure 4.1 présente les trajectoires balle/raquette pour une des simulations réalisée dans la session, et le portrait de phase du mouvement de la raquette lors de cette simulation. Les points noirs sur le portrait de phase représentent les impacts de la balle sur la raquette.

L'évaluation d'une simulation est réalisée grâce à la contrainte physique sur l'accélération de la raquette au moment de l'impact exprimée par la relation (2.1). Pour $g = 9.81\,m.s^{-2}$ et $\alpha = 0.48$, on doit avoir : -11.02 $< a_r <$ 0. D'après le tableau 4.1, l'accélération de la raquette au moment l'impact moyenne pour cette session est de -2.15. On a donc un rebond purement passif respectant la contrainte physique de la tâche.



(a) Trajectoires (b) Portrait de phase

Figure 4.1 – Rebond passif sans perturbation de l'environnement pour les conditions de référence $g = 9.81\,m.s^{-2}$ et $\alpha = 0.48$. Les points noirs sur les portraits de phase représentent les impacts de la balle sur la raquette.

	$g = 9.81\,m.s^{-2}$ et $\alpha = 0.48$
Vitesse de la raquette à l'impact v_r $(m.s^{-1})$	0.88 (0.0705)
Accélération de la raquette a_r $(m.s^{-2})$	-2.15 (2.605)
Phase de la raquette ϕ_r (deg)	267.77 (9.25)
Période de la raquette P_r (s)	0.66 (0.0)
Amplitude de la raquette A_r (m)	0.18 (0.0)

Table 4.1 – Données relatives à la raquette au moment de l'impact pour les conditions de référence $g = 9.81\,m.s^{-2}$ et $\alpha = 0.48$ (l'écart type pour chaque variable est donné entre parenthèses).

Variation de gravité

Les conditions initiales des systèmes dynamiques, présentées à la section 3.1, ont été déterminées pour les conditions environnementales de référence, dont un exemple a été montré ci-dessus. S'il est possible de re-calibrer les systèmes pour chaque condition environnementale différente, l'idée ici est de voir dans quelle mesure ces conditions initiales sont valides selon les conditions environnementales. Cinq sessions ont donc été réalisées, pour les valeurs de g suivantes : $6.56\,m.s^{-2}$, $8.10\,m.s^{-2}$, $9.81\,m.s^{-2}$, $11.66\,m.s^{-2}$, $13.69\,m.s^{-2}$.

Seules deux sessions ont permis d'observer un rebond maintenu : $g = 8.10\,m.s^{-2}$ et $g = 9.81\,m.s^{-2}$. Le système est donc très limité en terme de possibilité de conditions environnementales initiales, puisque seule une gravité initiale comprise entre 8.10 et $9.81\,m.s^{-2}$ permet le maintien du rebond. Pour ces deux valeurs, l'accélération de la raquette au moment de l'impact

vaut respectivement -3.47 et -2.01. Grâce à la relation (2.1), on sait que l'on a, pour des gravité de 8.10 et 9.81, les intervalles respectifs suivants : -9.09 $< a_r < 0$ et -11.02 $< a_r < 0$. Le rebond est donc non seulement maintenu, mais la contrainte physique caractéristique de la stabilité passive de la tâche est elle aussi respectée.

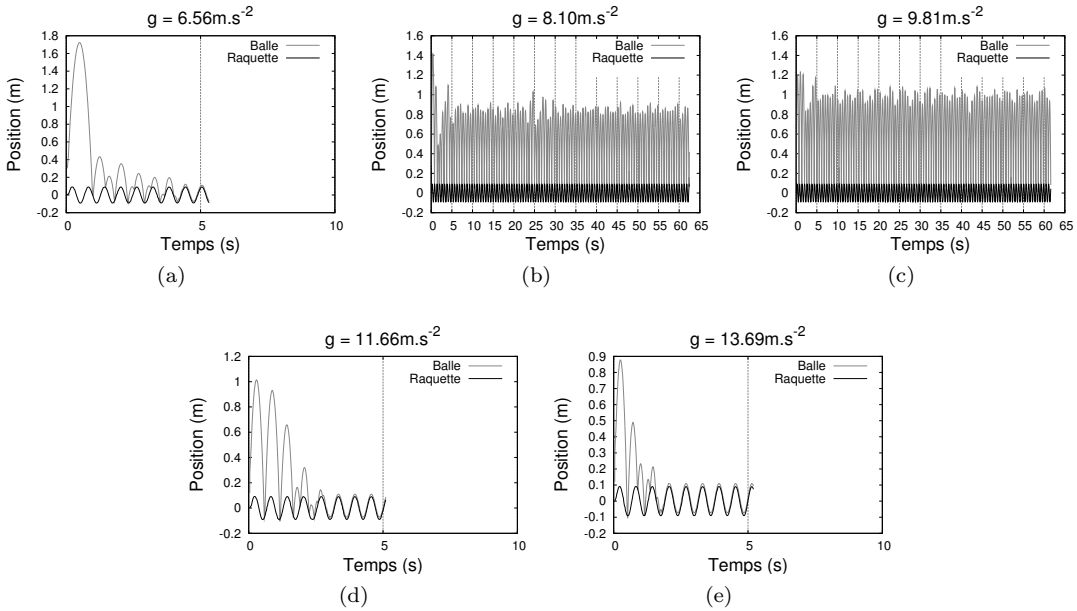


Figure 4.2 – Trajectoires pour différentes valeurs de g. Seules les valeurs $8.10\,m.s^{-2}$ et $9.81\,m.s^{-2}$ permettent de maintenir un rebond passivement stable.

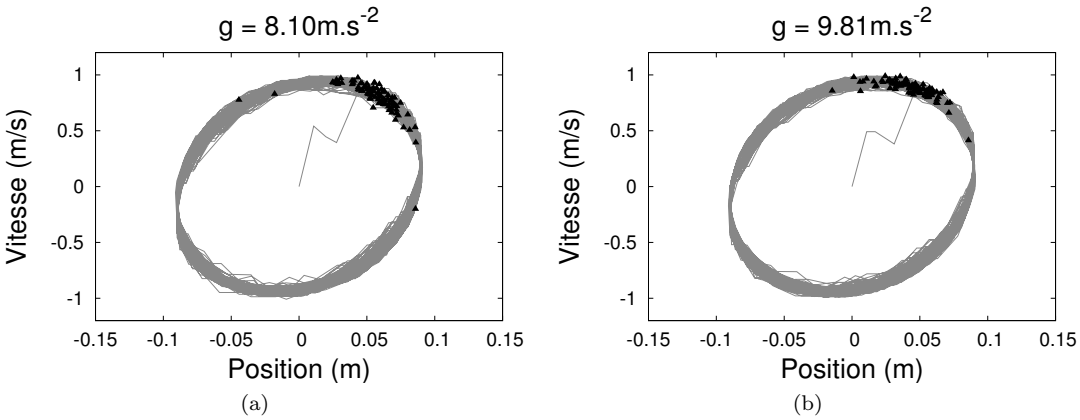


Figure 4.3 – Portraits de phase pour les deux valeurs de g permettant un rebond stable.

	$g = 8.10\,m.s^{-2}$	$g = 9.81\,m.s^{-2}$
v_r $(m.s^{-1})$	0.80 (0.102)	0.87 (0.063)
a_r $(m.s^{-2})$	-3.47 (3.47)	-2.01 (2.39)
ϕ_r (deg)	280.9 (9.64)	271.22 (9.33)
P_r (s)	0.66 (0.0)	0.66 (0.0)
A_r (m)	0.18 (0.0	0.18 (0.0)

Table 4.2 – Données relatives à la raquette au moment de l'impact pour différentes conditions environnementales initiales, constantes durant toute la durée d'une simulation (l'écart type pour chaque variable est donné entre parenthèses).

4.1.3 Conditions environnementales variables au cours d'une même simulation

Deux séries de simulations sont présentées ici. Les conditions environnementales sont modifiées au cours d'une même simulation afin de tester la capacité du système à compenser ces perturbations malgré l'absence de contrôle actif.

Dans la première série de simulations, la perturbation consiste en une diminution soudaine de la gravité : de $g = 9.81$, on passera à respectivement 4.0, 6.56 et $8.10\,m.s^{-2}$, correspondant donc à trois sessions en tout.

La seconde série de simulations est équivalente, à la différence que la perturbation est une augmentation soudaine de la gravité : de $g = 9.81$, on passe à respectivement 11.0, 11.66 et $12.3\,m.s^{-2}$, représentant également trois sessions.

Pour chaque simulation réalisée, la perturbation a lieu à $t = 15\,s$, lorsque la balle atteint sa hauteur maximale h_b.

Diminution soudaine de la gravité

La figure 4.4 présente des exemples de simulations réalisées pour chaque session de la série. Pour chacune des sessions, le rebond a pu être maintenu malgré l'introduction de la perturbation. Plus la baisse de la gravité est importante et plus la hauteur du rebond diminue. On note également, sur les portraits de phase, la disparité particulièrement grande des impacts pour $g = 4.0\,m.s^{-2}$.

Lorsque la gravité passe à 4.0, 6.56 et 8.10, l'accélération de la raquette à l'impact devrait théoriquement se trouver dans les intervalles respectifs : -4.49 $< a_r <$ 0, -7.37 $< a_r <$ 0 et -9.09 $< a_r <$ 0. Les accélérations respectives sont -5.43, -4.69 et -3.64 (tableau 4.3). Le rebond est donc stable passivement.

	$g = 9.81\,m.s^{-2}$ $\rightarrow 4.0\,m.s^{-2}$	$g = 9.81\,m.s^{-2}$ $\rightarrow 6.56\,m.s^{-2}$	$g = 9.81\,m.s^{-2}$ $\rightarrow 8.10\,m.s^{-2}$
v_r $(m.s^{-1})$	0.52 (0.41)	0.76 (0.128)	0.81 (0.101)
a_r $(m.s^{-2})$	-5.43 (4.30)	-4.69 (2.44)	-3.64 (2.56)
ϕ_r (deg)	284.8 (21.184)	281.69 (11.18)	278.19 (10.26)
P_r (s)	0.66 (0.0)	0.66 (0.0)	0.66 (0.0)
A_r (m)	0.18 (0.0	0.18 (0.0)	0.18 (0.0)

Table 4.3 – Données relatives à la raquette au moment de l'impact pour des simulations dans lesquelles une modification environnementale intervient, ici une diminution de la gravité (l'écart type pour chaque variable est donné entre parenthèses).

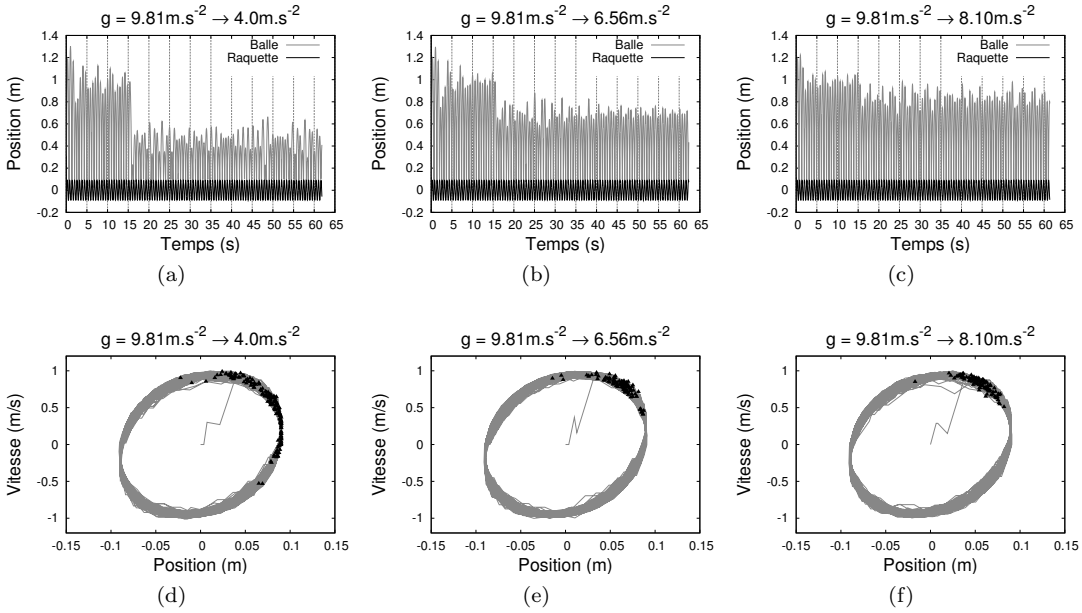


Figure 4.4 – Résultats de simulations pour des conditions environnementales variant au cours d'un même essai : diminution de la gravité.

Augmentation soudaine de la gravité

Des exemples de simulations réalisées lors de chacune des séries sont visibles sur la figure 4.5. Le rebond n'a pu être maintenu durant toute la durée de la simulation que dans le cas où $g = 11$, pour une accélération de la raquette à l'impact moyenne de -0.73. Suite à la perturbation, pour des valeurs de g supérieures à 11, le rebond ne peut être maintenu au delà d'un certain temps, qui est d'autant plus court que la gravité est haute (figures 4.5b et 4.5c).

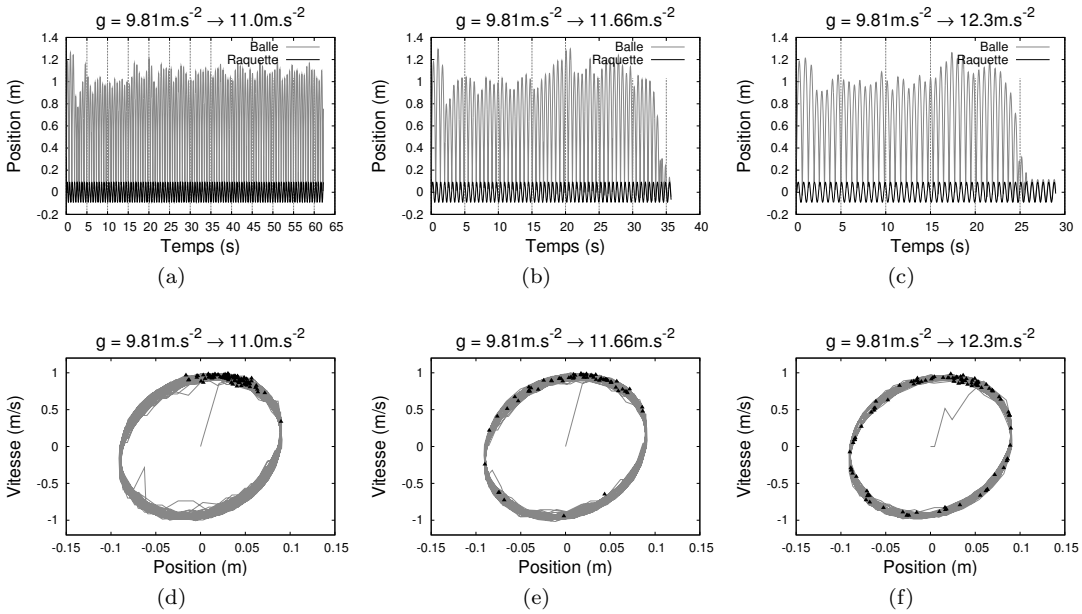


Figure 4.5 – Résultats de simulations pour des conditions environnementales variant au cours d'un même essai : augmentation de la gravité.

	$g = 9.81\, m.s^{-2} \rightarrow 11.0\, m.s^{-2}$
v_r $(m.s^{-1})$	0.91 (0.06)
a_r $(m.s^{-2})$	-0.73 (3.03)
ϕ_r (deg)	260.23 (11.05)
P_r (s)	0.66 (0.0)
A_r (m)	0.18 (0.0)

Table 4.4 – Données relatives à la raquette au moment de l'impact pour des simulations dans lesquelles une augmentation de la gravité intervient (l'écart type pour chaque variable est donné entre parenthèses).

4.1.4 Synthèse

Ces simulations ont permis de mettre en évidence le caractère passivement stable du rebond de balle. En outre, on a pu observer une certaine robustesse face aux différentes conditions environnementales utilisées. Lorsque la gravité est initialement basse, le système ne parvient pas à se stabiliser ; de même lorsque la gravité est initialement haute. La seule plage ayant permis l'observations de rebond maintenu est entre g = 8.10 et 9.81 $m.s^{-2}$.

L'introduction de la perturbation en cours de simulation est supportée par le système dans une certaine mesure, en particulier lorsque la gravité diminue. On note toutefois, pour toutes les variables de la raquette observées, un écart type fort pour une perturbation d'amplitude forte, lorsque g passe de 9.81 à 4.0 $m.s^{-2}$. Cela semble dénoter la difficulté du système à se stabiliser lors de la simulation. En revanche, une augmentation de gravité supérieur à 11.0 $m.s^{-2}$ ne permet pas de maintenir le rebond sur une longue durée.

Le système peut donc compenser des perturbations, mais ces perturbations doivent avoir une amplitude modérée. Il convient alors d'ajouter un contrôle actif, qui permettra de mieux gérer les perturbations de l'environnement et qui tire parti de la stabilité passive de la tâche.

4.2 Rebond avec contrôle actif

Dans cette section, les simulations présentées sont réalisée avec un contrôle actif. Comme précisé précédemment, ce contrôle consiste en un couplage entre les systèmes de la balle et de la raquette.

4.2.1 Méthode

Dans le cas du contrôle actif, trois séries de simulations sont réalisées :

1. Des tests sont d'abord réalisés sur le paramètres λ permettant de moduler les ajustements d'amplitude de la raquette, afin d'observer la réaction du système suivant les valeurs de λ. Ces simulations sont réalisées pour les conditions environnementales de référence : $g = 9.81\, m.s^{-2}$ et $\alpha = 0.48$.
2. Les conditions environnementales sont constantes au cours d'une même simulation.
3. Une perturbation de l'environnement est introduite au cours d'une même simulation.

Chaque simulation dure 60 secondes et les 10 premières secondes de chaque simulations sont retirées pour les calculs de moyennes et d'écart types des variables ainsi que pour le tracé des portraits de phase. Ces 10 secondes correspondent à une phase d'« amorce » du système, qui est due à la prise en compte d'une hauteur cible pour la balle. Cette hauteur cible est différente de la hauteur obtenue avec le rebond passif, et comme les conditions initiales sont les mêmes dans le cas présent, le système va faire des ajustements pour atteindre cette hauteur. Les données qui nous intéressent sont celles obtenues une fois que le système a atteint cette hauteur cible.

Pour chaque condition environnementale, quatre sessions de cinq simulations sont réalisées pour la première série (une session pour chaque valeur de λ ; λ = 0.2, 0.4, 0.6 et 0.8). Deux sessions de cinq simulations sont réalisées pour les deux autres séries, dans lesquelles les valeurs de λ utilisées sont 0.2 et 0.8.

En outre, chacune des trois séries est décomposée en deux sous-séries, en fonction des conditions environnementales retenues pour le couplage : t_{up} et h_b dans un premier temps, t_{up}, T_b et h_b dans un second temps.

On rappelle que les informations t_{up} et T_b sont utilisées pour procéder au ajustements de la période de l'oscillateur ; h_b permet de moduler l'amplitude de l'oscillateur.

Les exemples de trajectoires présentées dans cette section comportent également l'erreur mesurée entre la hauteur cible à atteindre et la hauteur effective de la balle.

4.2.2 Caractérisation

Cette première série de simulations a pour objectif l'étude du système pour les différentes valeurs de λ testées. Dans chacune des simulations, la gravité vaut $9.81\, m.s^{-2}$.

Informations visuelles utilisées : t_{up} et h_b

La figure 4.6 présente des exemples de simulations pour chaque valeur de λ. On rappelle que si le contrôle tire parti de la stabilité passive de la tâche, alors on doit avoir une accélération de la raquette à l'impact respectant la relation (2.1). Ici, pour g = $9.81\, m.s^{-2}$, on doit avoir : -11.02 $< a_r <$ 0.

Les moyennes des variables liées à la raquette sont présentées dans le tableau 4.5. Seule la moyenne des simulations pour $\lambda = 4$ est positive : une accélération à l'impact positive est possible lors du contrôle actif (Morice et coll., 2007) ; par ailleurs le rebond a été maintenu durant les simulations. Cependant, sachant que le contrôle mixte doit tirer parti de la propriété de stabilité de la tâche, l'accélération de la raquette à l'impact pour $\lambda = 0.4$ dénote une certaine incohérence. Celle-ci semble difficile à expliquer, une raison peut être l'utilisation du moteur physique (un environnement bruité) qui a une forte influence sur la déroulement des simulations.

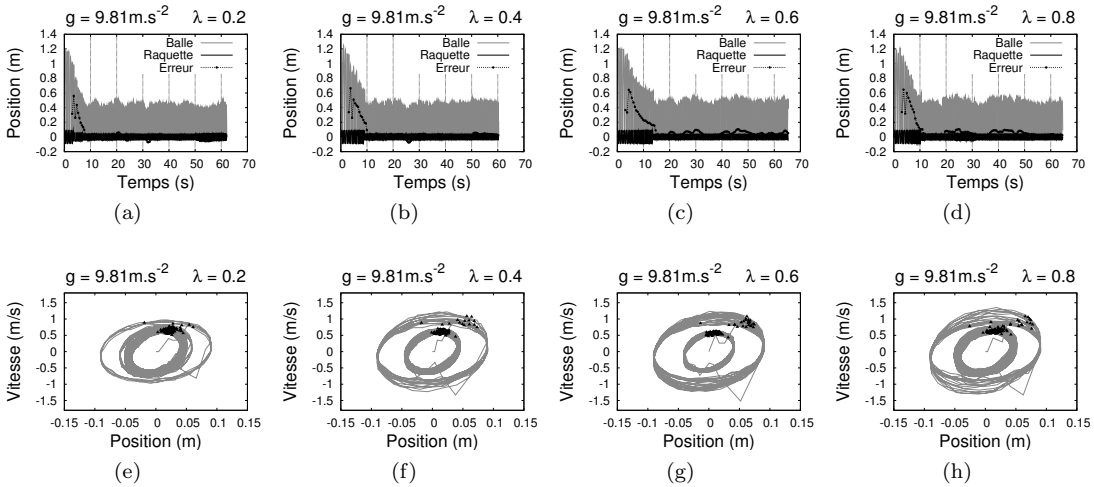


Figure 4.6 – Simulations en contrôle actif pour des conditions environnementales constantes au cours d'un même essai et pour $\lambda = 0.2,\ 0.4,\ 0.6,\ 0.8$. Les informations environnementales utilisées sont t_{up} et h_b.

	$g = 9.81\, m.s^{-2}$			
	$\lambda = 0.2$	$\lambda = 0.4$	$\lambda = 0.6$	$\lambda = 0.8$
v_r $(m.s^{-1})$	0.61 (0.28)	0.61 (0.058)	0.69 (0.133)	0.72 (0.16)
a_r $(m.s^{-2})$	-0.53 (4.81)	0.467 (2.30)	-0.88 (3.31)	-1.7 (3.88)
ϕ_r (deg)	276.53 (24.17)	269.94 (8.9)	275.77 (11.86)	278.76 (11.64)
P_r (s)	0.43 (0.04)	0.46 (0.023)	0.44 (0.027)	0.44 (0.03)
A_r (m)	0.09 (0.017)	0.086 (0.011)	0.097 (0.02)	0.103 (0.024)
Erreur ϵ (m)	0.071 (0.107)	0.039 (0.039)	0.04 (0.039)	0.045 (0.0467)

Table 4.5 – Moyennes des variables liées à la raquette pour des conditions environnementales constantes au cours d'une même simulation ($g = 9.81\, m.s^{-2}$) et pour $\lambda = 0.2,\ 0.4,\ 0.6,\ 0.8$. Les informations environnementales utilisées sont t_{up} et h_b.

Informations visuelles utilisées : t_{up}, T_b et h_b

Dans le cas présent, on prend en compte la totalité de la période de la balle pour ajuster la période de la raquette : un second ajustement est réalisé une fois que T_b a été perçu.

La figure 4.7 présente des exemples de simulations pour chacune des valeurs de λ. On observe plus de variabilité lors des simulations. Sur les portraits de phase, on note une plus grande disparité des points d'impacts. Le rebond a été maintenu lors de toutes les simulations. Dans le tableau 4.6, on note que les écarts type pour toutes les variables sont plus élevée du fait de la réalisation du second ajustement en fin de période de balle. En outre, l'erreur mesurée est plus importante que lorsqu'un seul ajustement de période est réalisé. Cependant, pour chaque valeur de λ, l'accélération de la raquette à l'impact est dans l'intervalle -11.02 $< a_r <$ 0, confirmant que le système a tiré parti la propriété de stabilité de la tâche.

Au vu des données présentées pour les deux sous-séries, il semble difficile de trancher sur les deux points suivants :

- La valeur de λ à utiliser pour ajuster l'amplitude suivant l'erreur perçue (toutes on permis le maintien d'un rebond sur l'ensemble des simulations).
- Les informations utilisées : qu'un seul ajustement soit fait sur la période de la raquette ou qu'un second ajustement soit réalisé en fin de période de balle, le rebond est maintenu dans tous les cas, et le contrôle tire parti de la propriété de stabilité de la tâche.

Pour ces raisons, dans les simulations suivantes, on procède de nouveau à deux sous-séries de simulations (une dans laquelle les informations environnementales utilisées sont t_up et h_b

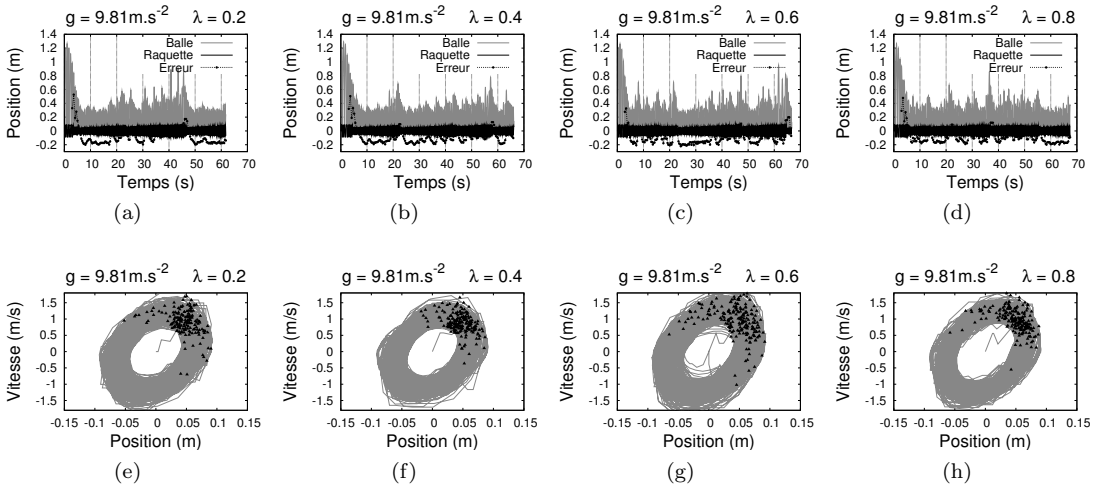


Figure 4.7 – Simulations en contrôle actif pour des conditions environnementales constantes au cours d'un même essai et pour $\lambda = 0.2,\ 0.4,\ 0.6,\ 0.8$. Les informations environnementales utilisées sont t_{up} et T_b et h_b.

	$g = 9.81\,m.s^{-2}$			
	$\lambda = 0.2$	$\lambda = 0.4$	$\lambda = 0.6$	$\lambda = 0.8$
v_r $(m.s^{-1})$	0.90 (0.36)	0.91 (0.37)	0.95 (0.45)	0.94 (0.43)
a_r $(m.s^{-2})$	-5.17 (8.54)	-4.98 (9.63)	-4.86 (11.66)	-4.69 (11.26)
ϕ_r (deg)	287.11 (21.55)	284.22 (22.78)	277.84 (26.36)	278.12 (25.05)
P_r (s)	0.36 (0.04)	0.36 (0.046)	0.366 (0.052)	0.36 (0.045)
A_r (m)	0.13 (0.023)	0.13 (0.023)	0.14 (0.027)	0.135 (0.024)
Erreur ϵ (m)	0.152 (0.078)	0.157 (0.08)	0.165 (0.089)	0.16 (0.086)

Table 4.6 – Moyennes des variables liées à la raquette pour des conditions environnementales constantes au cours d'une même simulation ($g = 9.81\,m.s^{-2}$) et pour $\lambda = 0.2,\ 0.4,\ 0.6,\ 0.8$. Les informations environnementales utilisées sont t_{up}, T_b et h_b.

uniquement ; l'autre dans laquelle les informations utilisées sont t_u, T_b et h_b).

Nous avons choisi d'utiliser les deux valeurs extrêmes de λ, 0.2 et 0.8, afin d'observer le comportement du système pour ces deux valeurs.

4.2.3 Conditions environnementales constantes au cours d'une même simulation

Dans cette seconde série de simulations, les conditions environnementales sont constantes au cours d'une même simulation. Le but est de tester des valeurs différentes de gravités afin d'observer le comportement du système : pour chaque valeur de λ, 0.2 et 0.8, deux sessions de simulations sont réalisées pour les valeurs de g suivantes : 6.56 et $13.69\,m.s^{-2}$. Ces deux valeurs de gravité ont été choisies afin de tester l'adabtabilité du système en contrôle actif. On rappelle que pour ces deux valeurs de g, le rebond passif n'a pu maintenir de rebond (paragraphe 4.1.2).

L'utilisation du mode de contrôle mixte doit permettre les observations suivantes (Siegler et coll., 2010) :

- Le système doit tirer parti de la propriété de stabilité passive de la tâche : l'accélération de la raquette à l'impact doit donc être compris dans l'intervalle défini à la relation (2.1).
  Pour $g = 6.56\,m.s^{-2}$, on a : $-7.37 < a_r < 0$.
  Pour $g = 13.69\,m.s^{-2}$, on a : $-15.38 < a_r < 0$.
- Une augmentation de g doit donner lieu à une diminution de la période (et inversement).

Informations visuelles utilisées : t_{up} et h_b

La figure 4.8 présente des exemples de simulations pour chaque valeur de g et de λ. Le système a plus de difficulé à se stabiliser lorsque g vaut $6.56\,m.s^{-2}$, ce qui est confirmé par les données présentées dans le tableau 4.7 : pour les deux valeurs de λ, les écarts type de chaque variable sont beaucoup plus élevée que lorsque g vaut $13.69\,m.s^{-2}$.

Toutefois, bien que pour une faible valeur de g la disparité des variables soit plus grande, le système tire parti de la propriété de stabilité passive de la tâche. En outre, on note effectivement une diminution de la période lorsque g vaut $13.69\,m.s^{-2}$.

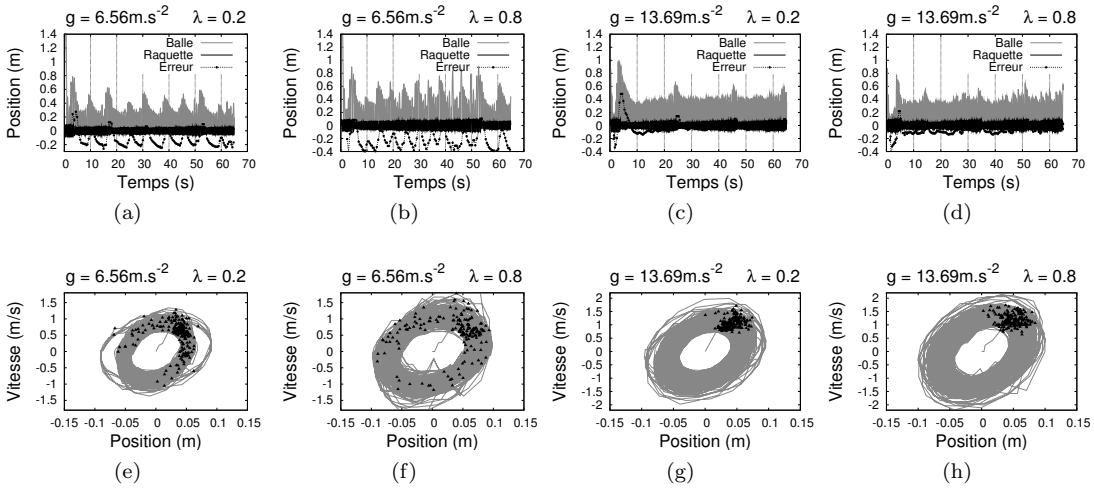


Figure 4.8 – Simulations en contrôle actif pour des conditions environnementales $g = 6.56\,m.s^{-2}$ et $g = 13.69\,m.s^{-2}$, constantes au cours d'un même essai et pour $\lambda = 0.2,\ 0.8$. Les informations environnementales utilisées sont t_{up} et h_b.

	$g = 6.56\,m.s^{-2}$		$g = 13.69\,m.s^{-2}$	
	$\lambda = 0.2$	$\lambda = 0.8$	$\lambda = 0.2$	$\lambda = 0.8$
v_r $(m.s^{-1})$	0.55 (0.43)	0.56 (0.54)	1.16 (0.25)	1.13 (0.22)
a_r $(m.s^{-2})$	-5.62 (8.88)	-4.23 (9.88)	-5.32 (7.87)	-4.86 (7.89)
ϕ_r (deg)	276.74 (26.07)	273.97 (31.99)	279.89 (12.62)	278.74 (21.16)
P_r (s)	0.42 (0.06)	0.46 (0.08)	0.34 (0.025)	0.33 (0.02)
A_r (m)	0.11 (0.015)	0.14 (0.02)	0.132 (0.02)	0.13 (0.02)
Erreur ϵ (m)	0.19 (0.1)	0.34 (0.17)	0.08 (0.05)	0.08 (0.04)

Table 4.7 – Moyennes des variables liées à la raquette pour des conditions environnementales constantes au cours d'une même simulation ($g = 6.56\,m.s^{-2}$ et $g = 13.69\,m.s^{-2}$) et pour $\lambda = 0.2,\ 0.8$. Les informations environnementales utilisées sont t_{up} et h_b.

Informations visuelles utilisées : t_{up}, T_b et h_b

Un second ajustement de période de raquette est réalisé dans cette seconde sous-série de simulations. Des exemples de simulations sont visibles sur la figure 4.9 et les moyennes des variables liées à la raquettes sont exposées dans le tableau 4.7. La différence avec la précédente série de simulation, dans laquelle un seul ajustement de la période de la raquette est effectuée, est la plus grande variabilité des simulations : les écarts type de chaque variables sont supérieurs à ceux présentés dans le tableau 4.7 ; ceci s'expliquant par l'ajout d'un second ajustement.

On note là encore la plus grande difficulté du système à stabiliser le rebond pour une faible valeur de g.

Dans tous les cas, le système tire parti de la propriété de stabilité passive de la tâche et la diminution de période de la raquette lorsque la gravité augmente.

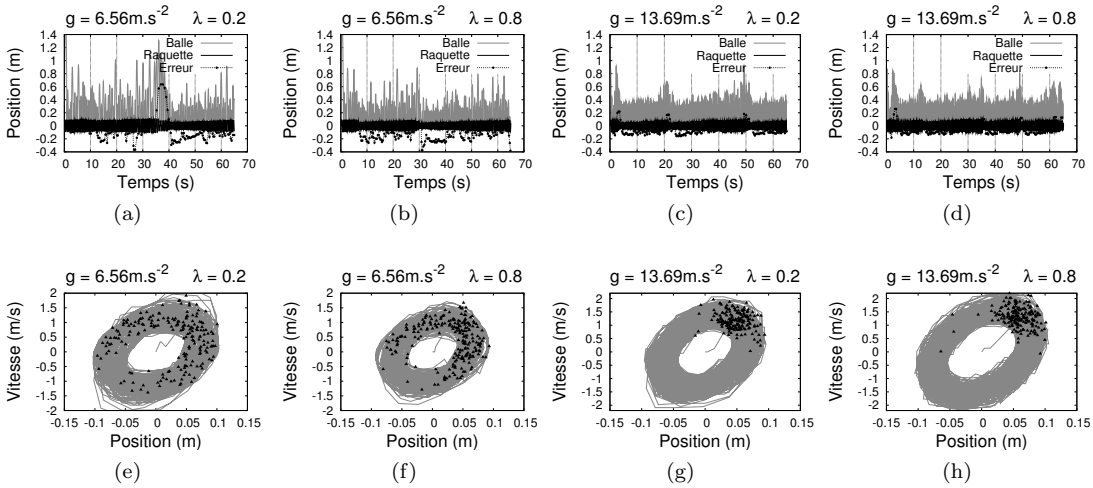


Figure 4.9 – Simulations en contrôle actif pour des conditions environnementales $g = 6.56\,m.s^{-2}$ et $g = 13.69\,m.s^{-2}$, constantes au cours d'un même essai et pour $\lambda = 0.2,\ 0.8$. Les informations environnementales utilisées sont t_{up}, T_b et h_b.

	$g = 6.56\,m.s^{-2}$		$g = 13.69\,m.s^{-2}$	
	$\lambda = 0.2$	$\lambda = 0.8$	$\lambda = 0.2$	$\lambda = 0.8$
v_r $(m.s^{-1})$	0.41 (0.74)	0.41 (0.66)	1.19 (0.32)	1.20 (0.34)
a_r $(m.s^{-2})$	-2.44 (13.47)	-3.01 (13.02)	-5.46 (10.10)	-5.81 (10.73)
ϕ_r (deg)	266.71 (36.36)	268.73 (36.21)	280.48 (16.71)	279.98 (18.64)
P_r (s)	0.44 (0.09)	0.43 (0.08)	0.33 (0.029)	0.33 (0.02)
A_r (m)	0.15 (0.03)	0.16 (0.02)	0.14 (0.027)	0.14 (0.03)
Erreur ϵ (m)	0.27 (0.14)	0.25 (0.13)	0.11 (0.06)	0.10 (0.05)

Table 4.8 – Moyennes des variables liées à la raquette pour des conditions environnementales constantes au cours d'une même simulation ($g = 6.56\,m.s^{-2}$ et $g = 13.69\,m.s^{-2}$) et pour $\lambda = 0.2,\ 0.8$. Les informations environnementales utilisées sont t_{up}, T_b et h_b.

4.2.4 Conditions environnementales variables au cours d'une même simulation

Dans cette dernière série de simulations, une perturbation est introduite dans l'environnement au cours d'une même simulation afin d'observer la réactivité du système. Ces perturbations sont de deux natures :

- Une baisse de gravité : de $g = 9.81\, m.s^{-2}$ on passe à $g = 6.56\, m.s^{-2}$.
- Une augmentation de gravite : de $g = 9.81\, m.s^{-2}$ on passe à $g = 13.69\, m.s^{-2}$.

Ces variations de gravité sont considérées comme d'amplitude importantes par Siegler et coll. (2010). Si dans le cas du rebond passif, le système ne pouvait supporter que des perturbations d'amplitude modérée, le contrôle actif doit permettre une réactivité accrue quelque soit l'amplitude de la perturbation.

Lors des simulations, la perturbation est introduite à $t = 25\, s$.

Informations visuelles utilisées : t_{up} et h_b

Des exemples de simulations sont visibles sur la figure 4.10 et les moyennes des variables liées à la raquettes sont visibles dans le tableau 4.9.

Dans cette série de simulation on note également la plus grande variabilité des rebonds lorsque g vaut $6.56\, m.s^{-2}$.

Le rebond est maintenu pour chaque perturbation et le système tire parti de la propriété de stabilité passive de la tâche.

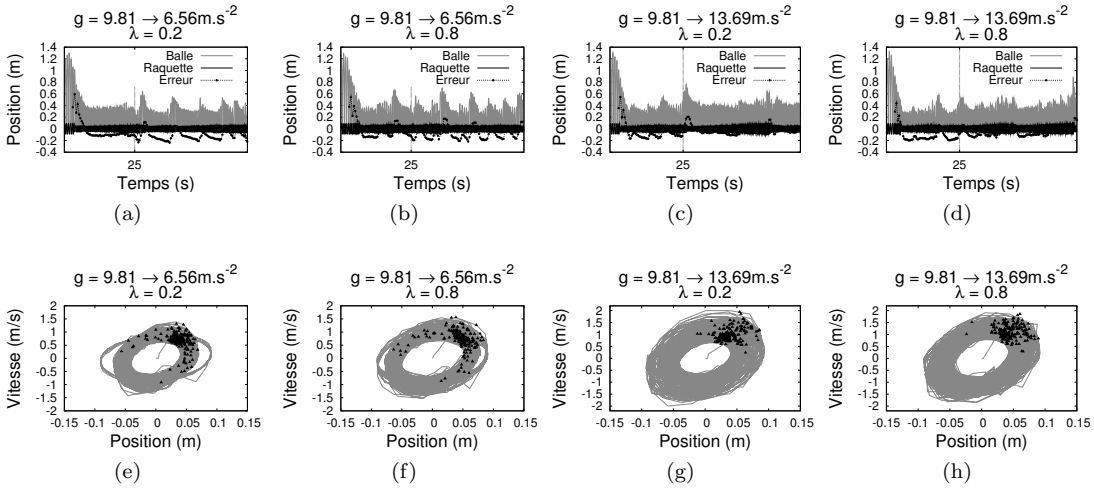


Figure 4.10 – Simulations en contrôle actif pour des conditions environnementales variant au cours d'un même essai (à $t = 25\, s$: $g = 6.56\, m.s^{-2}$ ou $g = 13.69\, m.s^{-2}$), et pour $\lambda = 0.2,\ 0.8$. Les informations environnementales utilisées sont t_{up} et h_b.

	$g = 9.81 \rightarrow 6.56\, m.s^{-2}$		$g = 9.81 \rightarrow 13.69\, m.s^{-2}$	
	$\lambda = 0.2$	$\lambda = 0.8$	$\lambda = 0.2$	$\lambda = 0.8$
v_r $(m.s^{-1})$	0.62 (0.40)	0.64 (0.40)	1.06 (0.28)	1.08 (0.26)
a_r $(m.s^{-2})$	-4.87 (7.80)	-5.42 (8.09)	-4.22 (8.03)	-4.64 (8.09)
ϕ_r (deg)	279.67 (23.49)	282.27 (22.87)	279.00 (13.47)	278.05 (14.73)
P_r (s)	0.41 (0.06)	0.41 (0.06)	0.35 (0.036)	0.35 (0.03)
A_r (m)	0.11 (0.010)	0.11 (0.02)	0.12 (0.02)	0.12 (0.02)
Erreur ϵ (m)	0.17 (0.09)	0.16 (0.09)	0.10 (0.06)	0.10 (0.06)

Table 4.9 – Moyennes des variables liées à la raquette pour des conditions environnementales variables au cours d'une même simulation ($g = 6.56\, m.s^{-2}$ et $g = 13.69\, m.s^{-2}$) et pour $\lambda = 0.2,\ 0.8$. Les informations environnementales utilisées sont t_{up} et h_b.

Informations visuelles utilisées : t_{up}, T_b et h_b

Dans cette seconde sous-série, un second ajustement de période est réalisé en fin de cycle de balle. Des exemple de simulations sont présentés sur la figure 4.11 et le tableau 4.10 contient les moyennes des variables liées à la raquette.

Pour cette série également, le rebond est maintenu pour les deux types de perturbation et la propriété de stabilité passive de la tâche est exploitée.

Le second ajustement de période entraîne une plus grande disparité des valeurs des variables de la raquette.

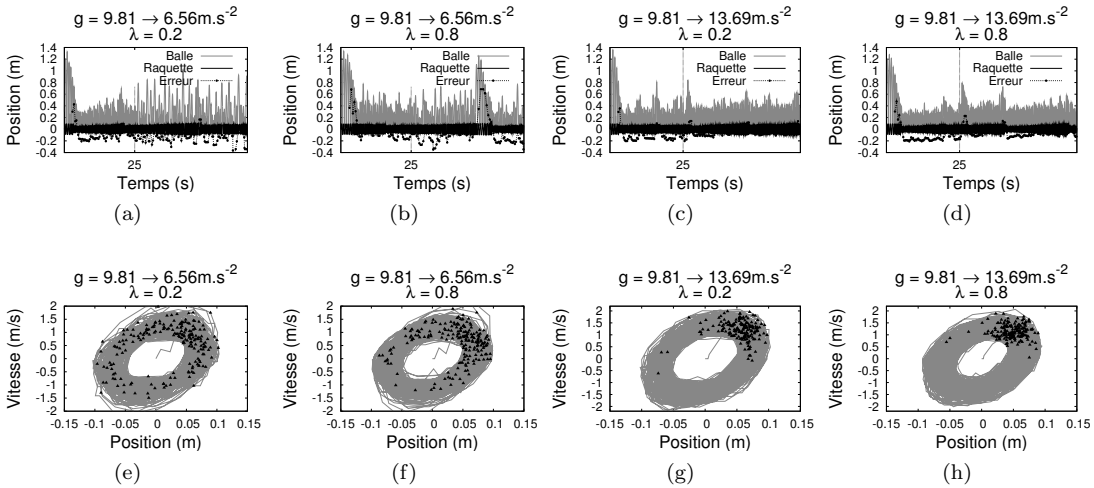


Figure 4.11 – Simulations en contrôle actif pour des conditions environnementales variant au cours d'un même essai (à $t = 25\, s$: $g = 6.56\, m.s^{-2}$ ou $g = 13.69\, m.s^{-2}$), et pour $\lambda = 0.2,\ 0.8$. Les informations environnementales utilisées sont t_{up}, T_b et h_b.

	$g = 9.81 \rightarrow 6.56\, m.s^{-2}$		$g = 9.81 \rightarrow 13.69\, m.s^{-2}$	
	$\lambda = 0.2$	$\lambda = 0.8$	$\lambda = 0.2$	$\lambda = 0.8$
v_r $(m.s^{-1})$	0.55 (0.73)	0.55 (0.67)	1.16 (0.41)	1.11 (0.33)
a_r $(m.s^{-2})$	-3.06 (13.83)	-3.81 (12.63)	-5.90 (11.71)	-5.37 (10.28)
ϕ_r (deg)	273.19 (31.194)	267.64 (35.5)	279.43 (23.26)	280.71 (19.17)
P_r (s)	0.41 (0.08)	0.41 (0.08)	0.34 (0.04)	0.33 (0.03)
A_r (m)	0.14 (0.03)	0.14 (0.02)	0.14 (0.03)	0.13 (0.02)
Erreur ϵ (m)	0.24 (0.13)	0.24 (0.12)	0.12 (0.07)	0.12 (0.07)

Table 4.10 – Moyennes des variables liées à la raquette pour des conditions environnementales variables au cours d'une même simulation (à $t = 25\, s$: $g = 6.56\, m.s^{-2}$ et $g = 13.69\, m.s^{-2}$) et pour $\lambda = 0.2,\ 0.8$. Les informations environnementales utilisées sont t_{up}, T_b et h_b.

4.2.5 Synthèse

Ces simulations ont permis de mettre en évidence la capacité du système à s'adapter non seulement à différentes conditions environnementales mais aussi face à des perturbations survenant dans l'environnement.

L'intérêt de l'approche par couplage de systèmes est de ne pas avoir décrit de comportement dans le système, mais d'observer l'émergence de ces comportements lors des simulations. En outre, le couplage tel qu'il est proposé par Siegler et coll. (2010) tire parti de la stabilité passive de la tâche, permettant de réduire l'amplitude des ajustements nécessaires, puisque les ajustements permettent non seulement de ramener le système dans son état stable mais aussi de l'y maintenir.

Cependant, les simulations proposées ici ne se veulent pas exhaustives. En effet, des choix ont du être effectués pour le couplage de certaines variables, notamment pour la modulation de l'amplitude de la raquette : la mesure de l'erreur au rebond permet d'ajuster cette amplitude, mais Siegler et coll. (2010) n'ont pu qu'émettre des hypothèses pour cet ajustement. Nous avons proposé d'utiliser la relation (3.4) pour mettre en place ces ajustements.

Nous avons testé différentes valeurs de λ pour ajuster l'amplitude de la raquette et proposé deux types d'ajustement pour la période de la raquette, mais il est impossible d'affirmer qu'une valeur ou un type d'ajustement est meilleur qu'un autre.

Toutefois, dans tous les cas, on observe l'émergence du comportement, et le contrôle actif mis en place est cohérent avec la définition du mode mixte proposé par Siegler et coll. (2010) : le contrôle est réalisé à chaque cycle et tire parti de la propriété de stabilité passive de la tâche.

Conclusion et perspectives

Dans le cadre de la dynamique de l'interaction entre entités autonomes et environnement, nous proposons d'utiliser une approche issue de la psychologie cognitive ecologique basée sur le couplage de systèmes dynamiques. Ce couplage doit donner lieu à une émergence de comportement dans le système et contourner les limites d'approches existantes à base de règles ou de raisonnement, qui peuvent présenter un manque de réactivité face à des perturbations de l'environnement et dont la connaissance du monde est restreinte. La possibilité d'obtenir des comportements réactifs existe dans ce type d'approche, grâce à la génération dynamique de règles permettant aux agents de réagir lorsque de nouveaux événements surviennent dans l'environnement (Bevacqua et coll., 2009).

Notre objectif est d'utiliser une approche qui ne nécessite pas d'expliciter les comportements, même dynamiquement. L'approche proposée par Warren (2006) est basée sur le couplage de systèmes dynamiques. Le couplage donne lieu à l'émergence de comportement dans le système, et le système se montre à la fois adaptatif et réactif. Nous avons mis en place un tel couplage grâce à une étude de cas, dont la tâche consiste à faire rebondir une balle sur une raquette. Différentes conditions environnementales sont utilisées lors des simulations et des perturbations sont introduites dans l'environnement. Le couplage permet au système de s'adapter à ces conditions, et de réagir face aux perturbations.

Compte tenu des approches existantes en terme d'animation, le travail proposé ici est somme toute modeste. Cependant, il ouvre une voie vers l'utilisation d'approches issues de la psychologie cognitive à base de systèmes dynamiques couplés ; l'objectif final étant de pouvoir étendre l'application d'une telle approches à différentes tâches implicants des interactions entre plusieurs entités voire avec des humains.

Au delà des limitations actuelles de l'approche utilisée ici (réalisation d'une tâche simple et très spécifique), le gain d'autonomie pour des entités interactives doit passer par l'obtention d'un comportement qui soit à la fois adaptatif, capable de surprendre tout en conservant une certaine régularité (De Loor et coll., 2009). L'utilisation de l'approche présentée dans ce travail présente de telles caractéristiques.

Références

Badler N.I. et Webber B.L., Planning and parallel transition networks : Animation's new frontiers, 1995.

Ballard D.H., « Animate vision », *Artif. Intell.*, vol. 48, p. 57–86, February 1991, ISSN 0004-3702, doi : 10.1016/0004-3702(91)90080-4.

Bevacqua E., Prepin K., de Sevin E., Niewiadomski R. et Pelachaud C., « Reactive behaviors in saiba architecture », Dans *Workshop Towards a Standard Markup Language for Embodied Dialogue Acts, held in conjunction with AAMAS'09, Budapest*, 2009.

Brooks R.A., « New approaches to robotics », *Science*, vol. 253, p. 1227–1232, 1991.

Buehler M., Koditschek D.E. et Kindlmann P.J., « Planning and control of robotic juggling and catching tasks », *The International Journal of Robotics Research*, vol. 13, n° 2, p. 101–118, 1994.

Cassell J., Pelachaud C., Badler N., Steedman M., Achorn B., Douville B., Prevost S. et Stone M., « Animated conversation : Rule-based generation of facial expression, gesture and spoken intonation for multiple conversational agents », Dans , p. 413–420, 1994.

De Loor P., Manac'h K. et Tisseau J., « Enaction-Based Artificial Intelligence : Toward Co-evolution with Humans in the Loop », *Minds and Machines*, vol. 19, n° 3, p. 319–343, août 2009.

Gibson J.J., *The Ecological Approach To Visual Perception*, Lawrence Erlbaum Associates, new edition édition, ISBN 0898599598, September 1986.

Keele S.W., Movement control in skilled motor performance, 1968.

Keijzer F.A., « Doing without representations which specify what to do », *Philosophical Psychology*, vol. 11, n° 3, p. 269–302, 1998.

Matsuoka K., « Sustained oscillations generated by mutually inhibiting neurons with adaptation », *Biological Cybernetics*, vol. 52, n° 6, p. 367–376, oct. 1985.

Matsuoka K., « Mechanisms of frequency and pattern control in the neural rhythm generators », *Biological Cybernetics*, vol. 56, n° 5, p. 345–353, juil. 1987.

Morice A.H.P., Siegler I.A., Bardy B.G. et Warren W.H., « Learning new perception-action solutions in virtual ball bouncing », *Experimental brain research*, vol. 181, p. 249–265, 2007.

Multon F., Ménardais S. et Arnaldi B., « Human motion coordination : a juggler as an example », *The Visual Computer*, vol. 17, n° 2, p. 91–105, 2001.

Prepin K. et Revel A., « Human-machine interaction as a model of machine-machine interaction : how to make machines interact as humans do », *Advanced Robotics*, vol. 21, n° 15, December 2007, adresse : `http://publi-etis.ensea.fr/2007/PR07`.

de Rugy A., Wei K., Müller H. et Sternad D., « Actively tracking 'passive' stability in a ball bouncing task », *Brain Research*, vol. 982, n° 1, p. 64 – 78, 2003.

Schaal S., Sternad D. et Atkeson C.G., « One-handed juggling : A dynamical approach to a rhythmic movement task », *Journal of Motor Behavior*, p. 165–183, 1996.

Siegler I.A., Bardy B.G. et Warren W.H., « Passive vs. active control of rhythmic ball bouncing : The role of visual information », *American Psychological Association : Human Perception and Performance*, vol. 36, p. 729–750, 2010.

Sternad D., Duarte M., Katsumata H. et Schaal S., « Bouncing a ball : Tuning into dynamic stability », *J. Exp. Psychol. Human Percept. Perform*, , n° 27, p. 1163–1184, 2001.

Todorov E. et Jordan M.I., « Optimal feedback control as a theory of motor coordination. », *Nature neuroscience*, vol. 5, n° 11, p. 1226–1235, November 2002, ISSN 1097-6256, doi : 10.1038/nn963.

Turvey M.T., Saltzman E. et Schmidt R.C., *Dynamics and task-specific coordinations*, p. 157–170, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, ISBN 1-55860-106-6, 1991.

Warren W.H., « The dynamics of perception and action », *Psychological Review*, vol. 113, n° 2, p. 358 – 389, 2006, ISSN 0033-295X.