



Exploration dynamique d'espaces multidimensionnels des paramètres de modèles de systèmes complexes

Encadrement : Gilles ARDOUDEL, Christian ATTIOGBÉ
Equipe : Vérification pour l'Environnement et le logiciel (VELO), LS2N
Mots-clés : Modèle paramétrique, exploration de modèles, simulation, modèle checking
Gratification : standard

1 Contexte et motivation

Ce travail se situe dans le cadre général de la modélisation et de l'analyse formelles de systèmes logiciels ou naturels. Nous entendons ici par système un ensemble d'éléments ou de composants interagissants. Nous avons entrepris de construire une plateforme logicielle (**Plug&Check**) qui, étant donné en entrée un modèle d'un système, permettra de l'analyser, par exemple avec des techniques d'exploration des états du système (*[statistical] model checking* [3]).

Pour des raisons d'efficacité, nous étudions les systèmes à travers des modèles paramétriques, gagnant ainsi en généralité et en couverture de familles de systèmes. Cependant, en plus de la question de la difficulté de construction des modèles, il y a la question de trouver les paramètres adéquats et les espaces de valeurs adéquats pour affiner les modèles et leur étude. Imaginons par exemple l'étude de mammifères marins ; ces animaux partagent de nombreux points communs mais ils sont en réalité très divers. En partant d'un modèle paramétrique des mammifères et en en faisant varier les paramètres, on peut alors envisager couvrir l'étude de plusieurs familles de ces animaux. Nous avons effectué des travaux préalables [1, 2] sur le sujet et continuons d'y travailler.

L'étude de l'ensemble des combinaisons possibles de valeurs de paramètres peut se faire en les discrétisant en une "grille" de valeurs. Pour chaque "case" de la grille, on observe alors le comportement du modèle et on en tire des conclusions sur le modèle et/ou les paramètres. Cette approche, bien que satisfaisante d'un point de vue de la couverture de l'ensemble des possibilités, peut s'avérer trop coûteuse si la grille est trop détaillée. De nombreuses questions se posent alors afin de réduire le coût effectif de cette approche.

On peut abstraire ce problème d'exploration de paramètres avec l'étude d'une grille de points (ou espace de points). Ces points peuvent aussi être des zones ou cellules (contenant une infinité de points). En fonction de la granularité choisie (par un effet de "zoom"), on étudie chaque zone de paramètres de manière plus ou moins détaillée. En fonction des résultats obtenus sur une zone, il peut être intéressant de "zoomer", i.e. découper cette zone en zones plus petites que l'on étudiera indépendamment.

L'échantillonnage de la grille et l'étude du comportement du modèle peuvent se paralléliser (sur plusieurs processeurs ou plusieurs machines) pour gagner en temps, mais cela soulève d'autres questions de synchronisation, ou de répartition des charges car certaines cellules peuvent provoquer des calculs détaillés et plus long alors que d'autres vont être simplement exclues.

2 Problématique

Parmi les questions et les problèmes sur lesquels nous travaillons, nous proposons d'aborder les suivants dans le cadre de ce stage de Master 2.

- **La définition des zones à explorer.** Nous cherchons des *méthodes précises* de découpage de l'espace des valeurs des paramètres, et l'impact des calculs locaux aux zones sur les autres calculs



environnants. La granularité des zones n'est pas figée, les résultats d'analyse du modèle peuvent provoquer des sous-découpages en zones plus fines (changement de dimensions).

Aussi, à priori, la durée de calcul peut seulement être grossièrement estimée pour l'espace de valeurs ou pour une cellule, car peuvent apparaître des zones à étudier en profondeur, donc à délimiter soigneusement.

- **Caractérisation de méthodes de parcours de l'espace des paramètres.** Le nombre de paramètres et les espaces de valeurs engendrent une difficulté et une complexité de parcours garantissant une bonne couverture de tout l'espace possible.
- **Borne des explorations et explorations en profondeur des zones.** Les résultats d'analyse du modèle varient en fonction des zones explorées et de leurs redimensionnements fins ; inversement, ces redimensionnements des zones impactent les résultats d'analyse. De nombreuses questions se posent ici (seuils des redimensionnements ; impacts croisés ; transfert de propriétés entre les dimensions locale et globale, ...). Un compromis est à trouver entre l'impact des résultats d'analyse sur les redimensionnements des zones étudiées ; en effet, l'arrêt précoce ou tardif des explorations en profondeur pourrait mettre en cause l'acuité des valeurs des paramètres.

3 Objectifs du stage et travaux envisagés

Les objectifs visés dans le cadre du stage sont listés ci-après, ainsi que les directions des travaux envisagés. En fonction de l'avancement du stage et de l'appétence du/de la stagiaire, on approfondira certains points plus que d'autres.

- Proposition de méthodes précises de définition de zones.

L'analyse d'un modèle donné sur une zone peut déclencher un découpage de cette zone et une exploration approfondie de ses sous-zones. Ainsi, de façon générale, les zones de l'espace dans lesquelles l'analyse du modèle montre une grande variabilité des résultats vont occasionner des traitements plus approfondis (et coûteux).

Plusieurs hypothèses sont à étudier dans ce contexte : il s'agira de déterminer et d'étudier les indicateurs qui permettront de déterminer lorsqu'il est opportun de subdiviser une zone donnée. Ces indicateurs peuvent être liés à la qualité de la zone (i.e. à la moyenne des résultats obtenus lors de son analyse), mais aussi à la qualité de l'analyse elle-même (i.e. à la variance obtenue lors de l'analyse de la zone). Les résultats obtenus devront ensuite être liés à la sensibilité de l'analyse globale du modèle.

- Détermination de l'avancement global d'une analyse.

Le processus d'analyse d'un modèle et de son jeu complet de paramètres peut être très long. Il est alors important de fournir à l'utilisateur une estimation de la progression de cette analyse en temps réel, par exemple au travers d'une "barre de progression". L'estimation de cette progression peut s'avérer compliquée. Il s'agira ici d'estimer l'avancement courant en fonction du nombre de zones traitées, tout en prenant en compte les redimensionnements dynamiques possibles des zones non traitées, l'avancement des calculs dans les zones en cours de traitement (et de leurs potentielles subdivisions) et les ressources disponibles à un instant donné.

- Recherche de méthodes de parcours les plus intéressantes pour couvrir l'espace des paramètres.

Le parcours le plus classique consiste en l'énumération du produit cartésien des découpages de chacune des dimensions, mais d'autres parcours peuvent être intéressants : *i)* priorisation d'une



dimension ou d'une zone *ii*) exploration radiale depuis des points ou des zones d'intérêt *iii*) exploration par échantillonnage non redondant et ultimement complet (analogue à une image dont les détails se précisent au fur et à mesure du chargement).

La description de l'état d'avancement (notamment pour reprise des calculs/explorations) devient alors plus complexe - la dimension interdisant de lister exhaustivement les cellules déjà parcourues ou à parcourir à chaque nouvel ordre de calcul.

- Recherche de solution pour borner/limiter le redimensionnement des zones des paramètres.

La variabilité des résultats d'analyse en fonction des zones ou sous-zones étudiées soulève plusieurs questions. Nous chercherons à déterminer : *i*) quels seuils et méthodes utiliser pour déterminer la nécessité de redimensionner en profondeur une zone ; *ii*) comment exploiter les résultats d'analyse d'une cellule pour déterminer l'influence de chaque paramètre sur la variabilité et donc le découpage de la grille des paramètres.

4 Environnement de travail

Le stage se fera dans l'équipe VELO, dans le cadre du projet de plateforme Plug&Check. Le stage sera encadré par Gilles Ardourel et Christian Attiogbé, mais bénéficiera des compétences, de la bonne humeur des autres membres de l'équipe et de la superbe ambiance qui y règne.

References

- [1] Étienne André, Benoît Delahaye, and Paulin Fournier. Consistency in parametric interval probabilistic timed automata. *J. Log. Algebraic Methods Program.*, 110, 2020.
- [2] David Julien, Guillaume Cantin, and Benoît Delahaye. End-to-end statistical model checking for parametric ODE models. In Erika Ábrahám and Marco Paolieri, editors, *Quantitative Evaluation of Systems - 19th International Conference, QEST 2022, Warsaw, Poland, September 12-16, 2022, Proceedings*, volume 13479 of *Lecture Notes in Computer Science*, pages 85–106. Springer, 2022.
- [3] Axel Legay, Benoît Delahaye, and Saddek Bensalem. Statistical model checking: An overview. In *Proc. Runtime Verification - First International Conference, RV 2010, St. Julians, Malta, November 1-4, 2010. Proceedings*, volume 6418 of *Lecture Notes in Computer Science*, pages 122–135. Springer, 2010.