Sujets de thèse autour du modèle AltaRica

L'équipe MVTsi*

4 mai 1998

Table des matières

1 Le thème scientifique

De plus en plus de systèmes produits par l'industrie contiennent des dispositifs logiciels complexes assurant des fonctions de contrôle-commande ou de communication dans des environnements plus ou moins prévisibles. Le bon fonctionnement de ces logiciels souvent critiques doit être absolument garanti :

- par des spécifications formelles précises et validées
- par des conceptions architecturales vérifiées
- par un développement méthodique et contrôlé
- par des tests aussi exhaustifs que possible

Les enjeux

Longtemps confinées au monde académique, les recherches dans ce domaine ont fait preuve de leur pertinence dans le monde industriel. Elles sont sur le point d'être utilisées massivement dans plusieurs domaines industriels (transports terrestres, aéronautique, énergie). Les chercheurs doivent donc être en mesure, dans un délai très court, de proposer aux industriels des techniques et des méthodes adaptées à leurs besoins.

*Document accessible à l'adresse http ://dept-info.labri.u-bordeaux.fr/ griffaul (ancrage html dans la rubrique Recherche)

Les verrous

Pour pouvoir répondre de façon satisfaisante à cette demande naissante, les efforts doivent porter sur trois points essentiels :

– l'intégration des diverses approches

Les chercheurs ont proposé et étudié plusieurs modèles de systèmes et plusieurs techniques de vérification. Aucun ne peut à lui seul prétendre couvrir tous les aspects d'une application réelle.
L'intégration de ces modèles et de ces techniques, qui est beaucoup plus que leur juxtaposition, demande encore beaucoup de recherches.

– la réflexion méthodologique

Pour être utile, l'offre de techniques issues de la recherche doit être accompagnée de leur "mode d'emploi". Cette réflexion méthodologique s'appuie sur des expériences nombreuses conduites en collaboration avec des partenaires industriels.

– le développement d'outils

Il est clair que ces techniques sont inséparables des outils qui les implémentent. Si dans une première phase les outils peuvent se présenter à l'état de maquette pour valider des algorithmes ou des structures de données spécifiques, il est indispensable qu'ils puissent être portés au moins à l'état de prototypes. Ceci dépasse bien entendu le domaine de la recherche pure et rentre en partie dans le cadre du transfert de technologie qu'il faut absolument et rapidement développer dans ce domaine.

2 Le projet AltaRica

Le modèle AltaRica

Le modèle AltaRica est un modèle basé sur les automates à contrainte. C'est une extension du modèle des systèmes de transition et de synchronisation étudié depuis longtemps dans l'équipe MVTsi.
Dans le modèle AltaRica, une transition n'est pas simplement :

```
''NomEtat'' |- ''EtiquetteTransition -> ''NomEtat''
```

mais plus généralement

```
''ContrainteSurVariables'' |- ''Evénement -> ''AffectationVariables'' priorité ''Entier''
```

Ce modèle à été choisi pour diverses raisons :

– sa simplicité,
– son pouvoir d'expression : il contient tous les modèles utilisés en vérification, test et sûreté de fonctionnement,
– sa proximité du modèle Arnold-Nivat.

Le langage AltaRica

Le langage AltaRica doit permettre de décrire un système complexe en terme d'un automate à contrainte. Pour cela divers concepts sont introduits afin de faciliter la tâche de l'utilisateur.

La hiérarchie un automate à contrainte peut-être décrit comme un arbre d'automates à contrainte. Cela permet un découpage fonctionnel, et facilite grandement la compréhension du système.

Variables d'états/variables de flux Les premières permettent de décrire les comportements des sous-systèmes (les noeuds de l'arbre des automates), les secondes les échanges de données entre eux (les arcs de l'arbre). Plus important, la sémantique de l'évolution d'un système en fonction du temps n'est pas la même pour ces deux types de données.

La synchronisation les événements sont par défauts indépendants, c'est à dire que leurs occurrences sont asynchrones, mais il est possible de contraindre plusieurs événements afin qu'ils soient simultanés.

En fait, il existe deux langages AltaRica, l'un textuel et l'autre graphique, mais totalement équivalent.

Les outils du projet AltaRica

Le langage de description AltaRica se veut fédérateur pour un ensemble de problèmes liés à l'étude des systèmes critiques et/ou complexes. Cela sous-entend que l'on puisse à partir d'une description AltaRica d'un système :

- calculer son arbre de défaillances, afin d'en extraire les coupes minimales et autres résultats qui interresseront le fiabiliste ;
- calculer des scénarios aboutissant soit à un événement redouté, soit à une situation particulière, afin de pouvoir tester si le système réel réagit conformément à ses spécifications.
- montrer que des propriétés de fonctionnement souhaitées sont bien réalisées.
- ...

Deux possibilités : un outil unique ayant l'ensemble des fonctionnalités, ou bien, un outil dédié à la saisie du modèle, et des compilateurs vers des outils spécialisés. Pour de nombreuses raisons, seul le second choix est raisonnable. Le projet AltaRica s'appuie donc sur un ensemble (non limité) d'outils, en fait représentatif de modèles mathématiques dérivables à partir des automates à contraintes.

Modèle Booléen Aralia, outil de calcul sur les arbres de défaillances.

Réseau de Petri Arabica, outil de calcul sur les réseaux de Petri stochastiques.

Automates à contraintes Toupie, μ calcul sur les domaines finis.

Système de transition Mec, vérificateur de modèles.

...

Les objectifs

Le projet AltaRica, piloté par le LaBRI, est financé en partie par des industriels et par la région Aquitaine. De ce fait les objectifs de différentes natures sont définis régulièrement par les membres du club AltaRica. La phase 2 (l'actuelle) du projet comporte :

du cours terme Cela regroupe principalement des tâches de développement pour lesquelles les travaux de recherches ont déjà été effectués.

- une consolidation du langage de description AltaRica.
- le compilateur d'un arbre d'automates en un automate.
- le compilateur d'un automate vers les arbres de défaillances.
- le compilateur d'un automate vers les réseaux de Petri.
- l'interface de saisie d'une description AltaRica.
- la simulation d'une description AltaRica.

du moyen terme Cela regroupe principalement des travaux de recherches sur lesquels il est nécessaire de travailler maintenant, afin d'aboutir soit à la faisabilité, soit au rejet.

- la vérification à la volée d'une description AltaRica.
- l'introduction du temps dans le modèle AltaRica (cf sujet 1).
- les logiques (cf sujet 3) et les langages d'expression (cf sujet 1) des propriétés d'un système.
- le compilateur vers les vérificateurs de modèles (cf sujet 3).
- les vérifications de propriétés : preuve, symbolique (cf sujet 2) et modulaire (cf sujet 4).
- la génération de scénarios (cf sujet 4).

mais également :

- les méthodologies (cf sujet 1) de modélisation et de vérification.

du long terme Cela regroupe uniquement des travaux de recherches qui pourrait influencer les phases ultérieures du projet.

- les preuves formelles (cf sujet 5) afin de pouvoir traiter des systèmes plus importants.
- l'interprétation abstraite afin de diminuer la taille du problème par une simple analyse de la description.

3 Sujet 0 : Modélisation et vérification temporelles en AltaRica

Fiche signalétique

Responsable	Antoine Rauzy	05 56 84 60 83	Antoine.Rauzy@labri.u-bordeaux.fr
Collaborateur	Alain Griffault	05 56 84 69 15	Alain.Griffault@labri.u-bordeaux.fr
Equipe	MVTsi		
Titre	Modélisation et vérification temporelles en AltaRica.		
Catégorie	Théorique et expérimentale		
Position/AltaRica	Définition du langage		
	Court et Moyen termes		
Financement	Bourse de thèse MENRT		

Présentation du sujet

Contexte Les systèmes critiques et/ou complexes représentant des systèmes industriels font très souvent références au temps. Celui-ci sert pour indiquer des informations de natures quelques fois très différentes :

- Des informations statistiques et/ou stochastiques sur les occurrences de certains événements.
- Des informations de causalité entre événements.
- Des délais pendant lesquels des événements sont impossibles.
- Des délais pendant lesquels des événements sont possibles.
- Des durées pour l'accomplissement de certains événements.

Le sémantique de l'actuelle version du langage AltaRica vis à vis du temps est la suivante : les événements et le changement global d'état qui en découle sont de durée nulle. C'est à dire, que le langage AltaRica permet de décrire des systèmes réactifs, ou le temps est assimilable aux entiers naturels.

Le langage AltaRica semble donc insuffisant pour traiter un certain nombre de problèmes pour lesquels il a été défini.

Travail demandé Le sujet de la thèse est l'étude du temps dans le projet AltaRica. Un premier travail consistera à recenser et à formaliser les besoins en terme d'ordonnancement, d'observateurs et de contrôleur de modèle AltaRica.

Une seconde étape consistera à comparer les puissances d'expression des modèles d'automates temporisés et/ou hybrides par rapport aux automates à contraintes AltaRica.

La phase terminale devra, non seulement aboutir à des préconisations pour le projet AltaRica ; mais également à des recommandations d'ordre méthodologiques pour la description en AltaRica d'un système temporisé.

Environnement La thèse s'appuiera évidemment sur les exemples temps-réels académiques, mais également sur plusieurs exemples de systèmes industriels qui sont d'ores et déjà disponibles (au LaBRI) et font l'objet de collaboration entre le LaBRI et divers partenaires industriels (CEA, Aerospatiale).

4 Sujet 1 : Intégration des techniques de preuve de théorèmes et de vérification symbolique pour la preuve de circuits séquentiels

Fiche signalétique

Responsable	Antoine Rauzy	05 56 84 60 83	Antoine.Rauzy@labri.u-bordeaux.fr
Collaborateur	David Sherman	05 56 84 69 22	David.Sherman@labri.u-bordeaux.fr
Equipe	MVTsi		
Titre	Intégration des techniques de preuve de théorèmes et de vérification symbolique pour la preuve de circuits séquentiels.		
Catégorie	Théorique et expérimentale		
Position/AltaRica	Outil satellite		
	Moyen terme		
Financement	Bourse de thèse BDI ou MENRT		

Présentation du sujet

Contexte Le comportement de nombreux systèmes industriels peut être modélisé par un automate booléen avec entrée-sorties (une machine de Mealy). La vérification de propriétés sur ce type de modèles se fait jusqu'à présent par parcours de l'espace des états accessibles de l'automate. L'introduction de techniques de parcours symbolique (utilisant un codage par diagrammes binaires de décision, BDD) a permis des progrès considérables dans la taille des systèmes traitables. Toutefois, de nombreux systèmes réels restent hors de portée, en raison de l'explosion combinatoire du nombre d'états à explorer, ou de la taille des BDD codant ces états. Pour aller plus loin, une technique prometteuse consiste à décomposer les automates étudiés suivant la propriété à vérifier. Or, il faut prouver formellement que ce type de décomposition est correcte.

Travail demandé Le sujet de la thèse est donc d'étudier l'intégration d'assistants de preuve du type PVS et d'outils de vérification symbolique du type SMV, en partant des travaux déjà effectués sur ce thème.

Environnement La thèse s'appuiera sur plusieurs exemples de systèmes industriels qui sont d'ores et déjà disponibles (au LaBRI) et font l'objet de collaboration entre le LaBRI et divers partenaires industriels (CEA, Aerospatiale).

5 Sujet 2 : Extensions de structures de données partagées pour la vérification de systèmes complexes

Fiche signalétique

Responsable	André Arnold	05 56 84 60 94	Andre.Arnold@labri.u-bordeaux.fr
Collaborateurs	David Sherman	05 56 84 69 22	David.Sherman@labri.u-bordeaux.fr
	Jean-Michel Couvreur	05 56 84 69 34	Jean-Michel.Couvreur@labri.u-bordeaux.fr
Equipe	MVTsi		
Titre	Extensions de structures de données partagées pour la vérification de systèmes complexes.		
Catégorie	Théorique et expérimentale		
Position/AltaRica	Outil satellite		
	Moyen terme		
Financement	Bourse de thèse MENRT		

Présentation du sujet

Contexte Le vérification de systèmes complexes s'appuie de plus en plus sur des techniques et outils de caractère *mixte*, intégrant de différents formalismes et moyens de mise en oeuvre. De par les problèmes d'-explosion combinatoire engendrés par les systèmes complexes, ces différentes techniques doivent utiliser des structures de données représentant de façon très compacte les systèmes à étudier. Les diagrammes binaires de recherche (BDD) et arbres partagés en sont des exemples répandus.

Travail demandé Le sujet de la thèse est d'étudier et de mettre en œuvre sur ces structures partagées un certain nombre d'opérations élémentaires (ensemblistes, quantifications, points fixes) et de bâtir sur elles un vérificateur pour des logiques plus puissantes que les logiques temporelles. Une réflexion sur les relations entre le mu-calcul et la logique temporelle permettra l'étude du pouvoir d'expression de ces logiques.

Environnement La thèse s'appuiera sur les systèmes existants Mec et Toupie et leur intégration dans un nouvel outil de vérification basé sur le système Factotum. L'outil ainsi réalisé se trouvera en aval du projet AltaRica, et exploitera des techniques issus du projet Clovis.

6 Sujet 3 : Techniques de vérification et de génération de séquences de test

Fiche signalétique

Responsable	Richard Castanet	05 56 84 69 01	Richard.Castanet@labri.u-bordeaux.fr
Collaborateur	Pierre Castéran	05 56 84 69 31	Pierre.Casteran@labri.u-bordeaux.fr
Equipe	MVTsi		
Titre	Techniques de vérification et de génération de séquences de test.		
Catégorie	Théorique et expérimentale		
Position/AltaRica	Outil satellite		
	Moyen terme		
Financement	Bourse de thèse MENRT		
	Frais de fonctionnement sur projet FORMA		

Présentation du sujet

Contexte L'équipe MVTsi (Modélisation, vérification et test de systèmes informatisés) possède des compétences dans les domaines de la vérification, du test et de la sûreté de fonctionnement. Les techniques de preuve représentent une compétence au LaBRI venant conforter le domaine de recherche lié au sujet de thèse. Les techniques les plus récentes dans la génération de séquences de test utilisent des techniques liées à la vérification des systèmes. Une démarche connue des milieux de la recherche et de plus en plus acceptée par les industriels est la suivante : spécification formelle d'un système, vérification formelle et génération automatique de séquences de test qui seront injectées dans le système implémenté. Le contexte général du sujet de recherche s'inscrit dans un projet intitulé FORMA (CNRS-MENSR) en collaboration avec l'ENS-Cachan, l'Irisa de Rennes, Verimag de Grenoble et le CNET-Lannion. Ce sujet sera aussi lié au projet AltaRica du LaBRI.

Travail demandé A partir de descriptions effectuées dans des langages de spécification formelle (en général SDL ou AltaRica), il s'agit de vérifier ou de prouver des propriétés et de permettre des générations automatiques de séquences de test selon des objectifs de test définis.

On s'intéressera à des approches prenant en compte une spécification par modules et en étant soit descendantes : vérification ou preuves de propriétés globales et génération de test global avec vision des implications sur les modules internes soit descendantes : vérification ou preuve d'un module de la spécification et génération de tests locaux avec intégration dans un test global. Des approches mixtes sont souhaitables : aide à la vérification à l'aide de techniques de preuve, aide à la génération automatique de tests à l'aide de techniques de vérification et de preuve.

Il faudra concevoir un outil global permettant à partir de spécifications formelles (SDL, AltaRica ...) de spécifier des propriétés (locales ou globales), de faire des vérifications et de générer automatiquement des tests.

Environnement Utilisation des outils

- Logiciel Geode (spécification et vérification SDL)
- Logiciel Mec
- Atelier AltaRica
- COQ

afin de traiter des exemples de :

- Protocoles de communication
- Systèmes industriels et sûreté de fonctionnement

7 Sujet 4 : L(A)=L(B) ?

Fiche signalétique

Responsable	Paul-Yves Gloess	05 56 84 69 27	Paul.Gloess@labri.u-bordeaux.fr
Collaborateurs	Pierre Castéran	05 56 84 69 31	Pierre.Casteran@labri.u-bordeaux.fr
	Géraud Sénizergues	05 56 84 60 48	ges@labri.u-bordeaux.fr
Equipe	Programmation symbolique		
Titre	L(A)=L(B) ?		
Catégorie	Théorique		
Position/AltaRica	Outil satellite d'une version ultérieure		
	Prospectif		
Financement	Bourse de thèse MENRT		

Présentation du sujet

Objectifs L'objectif essentiel de la thèse est d'obtenir une version "mécanisée" de la preuve du théorème de Géraud Sénizergues : "L(A)=L(B) est décidable, A et B étant des automates à pile finis déterministes", sous une forme convaincante pour le mathématicien ou le théoricien de l'informatique. Il faut noter l'importance du résultat obtenu par Géraud Sénizergues, résultat qui fut pendant longtemps une conjecture intéressant la communauté scientifique internationale. Au delà de cet objectif, qui constitue en lui même un véritable "challenge", on s'efforcera d'abstraire les principes généraux de la preuve, sous la forme d'une bibliothèque de théories et de stratégies de preuve applicables à d'autres résultats de décidabilité.

Contexte La preuve rédigée par Géraud Sénizergues dans `~ges/public_html/lalb_l.ps` (version améliorée du rapport LaBRI RR-1161-97) compte 84 pages de texte, jusqu'au résultat principal, le Théorème 9.3, qui énonce la décidabilité du problème d'équivalence, pour des automates à pile finis déterministes. Le texte est dense, et les preuves des différents lemmes non triviales, mais il est relativement auto-suffisant : les objets mathématiques du domaine sont pour la plupart définis ; les théorèmes extérieurs au texte sont en petit nombre : ce sont des résultats classiques dont les preuves sont facilement accessibles dans les livres. Cet aspect, renforcé par la présence au LaBRI de l'auteur de la preuve, rend l'entreprise de mécanisation envisageable, au moyen d'assistants de preuves internationalement reconnus tels que PVS (SRI International) ou Coq (INRIA) disponibles au LaBRI.

Démarche et travail demandé Il ne s'agit pas uniquement d'obtenir une preuve du Théorème 9.3 au moyen de PVS ou Coq. Encore faut-il que le théoricien (tel que l'auteur de la preuve) et le praticien (utilisateur de PVS ou Coq) puissent être convaincus du bien-fondé de la preuve mécanique, et de l'identité du résultat mécaniquement démontré avec le Théorème 9.3. Cette exigence implique en particulier qu'aucun axiome ne soit introduit, pour pallier la preuve de résultats extérieurs utilisés, et que les définitions des notions sous-jacentes (par exemple, les automates à pile, la décidabilité) soient présentées sous une forme accessible et indiscutable au théoricien.

En pratique, on aura intérêt à suivre de près la preuve humaine, sous peine de la réinventer, ce qui pourrait prendre beaucoup plus de temps, et n'est pas demandé. D'autre part, dans une perspective de réutilisabilité des théories PVS ou Coq, et plus généralement d'intéressement du théoricien aux résultats de la thèse, il est impératif d'établir une correspondance limpide entre le texte initial et la version mécanisée.

Bien que le résultat principal soit un résultat de décidabilité, de nombreux lemmes utilisent des preuves non constructives, de sorte que le choix du meilleur outil n'est pas évident. L'impact des différences entre les logiques de PVS et Coq, sur la faisabilité de l'entreprise, devra être finement analysé avant de prendre une décision. Certaines de ces différences sont majeures : distinction existant en Coq (pas en PVS) entre les sortes Set et Prop ; notion d'égalité ; typage décidable en Coq et indécidable en PVS. Elles ont un impact sur le langage des fonctions récursives, et la nature des preuves. D'autres différences plus mineures peuvent néanmoins prendre de l'importance en pratique, étant donné l'ampleur du domaine à traiter : surcharge possible en PVS, pas en Coq ; currification implicite en Coq, identification des types [D1, ..., Dn $\to$ D] et [[D1, ..., Dn] $\to$ D] en PVS.

On comparera dans une première phase de quelques mois d'expérimentation les deux logiques sur un échantillon représentatif du texte de Géraud Sénizergues, incluant en particulier l'échantillon abordé par une étude préliminaire de faisabilité sous PVS (`~gloess/pvs/ges/README`). Le choix de l'outil reposera sur une confrontation des résultats de cette comparaison avec les objectifs énoncés plus haut. La confiance que l'on peut avoir dans l'outil lui-même est liée à la conception de celui-ci : c'est un aspect à ne pas négliger.

On achèvera dans une seconde phase la formalisation et les preuves, jusqu'au Théorème 9.3, à l'aide de l'outil choisi.

La troisième phase visera l'objectif de réutilisabilité.