

SORBONNE UNIVERSITÉ

RAPPORT FINAL

19 Mai 2023

---

# Projet STL : Interface pour Autobill

---

*Auteurs :*

Brahima DIBASSI  
Zeid FAZAZI  
Yukai LUO

*Encadrants :*

Hector SUZANNE  
Emmanuel CHAILLOUX



**SORBONNE  
UNIVERSITÉ**

# Table des matières

<b>1</b>	<b>Contexte du projet</b>	<b>2</b>
1.1	Historique et définitions . . . . .	2
1.2	Qu'est-ce qu'Autobill? . . . . .	3
1.3	Objectifs du projet . . . . .	4
1.4	Processus de conception . . . . .	4
<b>2</b>	<b>Interface web</b>	<b>5</b>
2.1	Client . . . . .	5
2.1.1	Interopérabilité Web-Ocaml . . . . .	5
2.1.2	L'éditeur de code . . . . .	7
2.1.3	La résolution de contraintes . . . . .	9
2.2	Serveur + client . . . . .	10
2.2.1	MiniML . . . . .	10
2.2.2	La résolution de contraintes . . . . .	12
<b>3</b>	<b>MiniML</b>	<b>13</b>
3.1	Description de MiniML . . . . .	13
3.2	Grammaire . . . . .	13
3.3	Call-By-Push-Value . . . . .	15
3.4	Semantique de traduction . . . . .	15
<b>4</b>	<b>Conclusion</b>	<b>20</b>
	<b>Références</b>	<b>21</b>

# 1 Contexte du projet

Dans le cadre de sa thèse sur l'analyse statique de consommation mémoire d'un programme, notre tuteur de projet, Hector Suzanne, membre de l'équipe APR du laboratoire LIP6, a développé [1].

L'analyse statique est un ensemble de méthodes formelles permettant de mesurer et détecter automatiquement des comportements ou des erreurs dans un programme en examinant son code source. Parmi les usages les plus courants de l'analyse statique, il y a le débogage pour identifier des erreurs syntaxiques (fautes de frappe) ou l'usage de variables non déclarées ou non-initialisées. D'autres outils emploient des heuristiques et des règles pour répondre à des problématiques d'optimisation de code ou de fiabilité/sécurité. Ces analyseurs statiques sont conçus dans le but d'améliorer la maintenabilité du code.

Autobill s'intéresse particulièrement à l'occupation mémoire d'un programme. Historiquement, ce sujet de recherche a été plusieurs fois abordé dans divers travaux scientifiques, parmi eux, ceux de Jan Hoffmann et Steffen Jost sur l'analyse de consommation de ressources automatisé AARA [2] (*Automatic Amortized Resource Analysis*).

Cette technique suit un ensemble de règles d'inférences : chaque trait du langage analysé est annoté par ses bornes mémoire et on peut déduire la consommation mémoire que ce trait va engendrer. Pour déterminer ces bornes, l'AARA s'aide de l'analyse amortie par la méthode du potentiel. En effet, l'analyse amortie exploite la structure séquentielle des opérations sur des structures de données dynamiques pour calculer le coût total de cette séquence. Avec la méthode du potentiel, on attribue un potentiel aux structures de données manipulées, ce qui va affecter le coût que va avoir une opération sur cette structure. L'application de ces règles d'inférence sur l'ensemble des expressions du programme fournit une estimation correcte de ses bornes mémoires.

## 1.1 Historique et définitions

Des langages de programmation expérimentaux ont vu le jour et ont implémenté cette analyse comme Resource Aware ML [3], un langage fonctionnel à la ML créé par Jan Hoffmann, Klaus

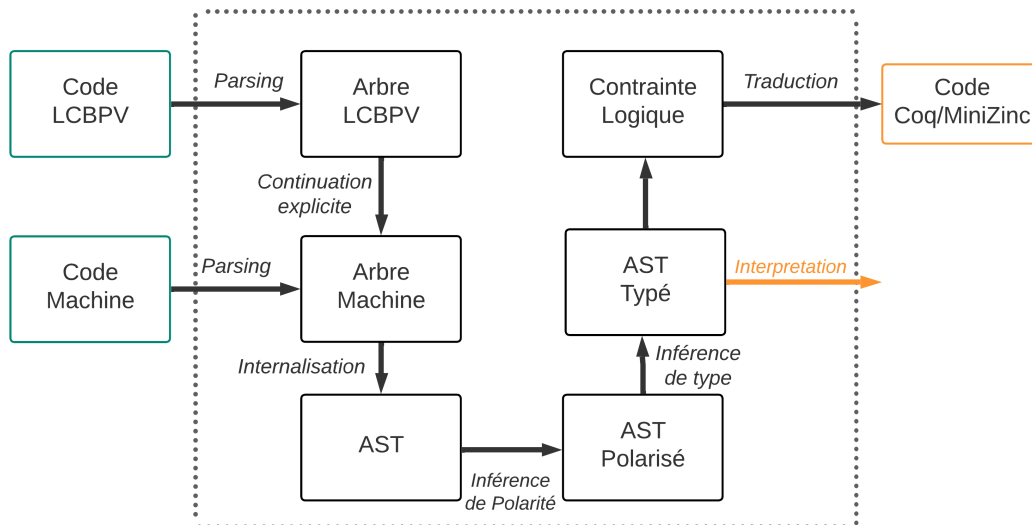


FIGURE 1 – Représentation simplifiée d'Autobill

## 1.2 Qu'est-ce qu'Autobill ?

La proposition d'Hector Suzanne avec Autobill se différencie par un niveau d'analyse plus précis sur les fermetures et les arguments fonctionnels d'un programme par rapport à RAML. D'abord, parmi les entrées possibles illustrées sur la gauche de la figure 1, Autobill ne supporte uniquement que des programmes écrits soit en modèle machine propre à Autobill, soit en **Call-By-Push-Value** (CBPV) [4], avec ou sans continuation explicite.

Call-By-Push-Value est un langage qui utilise un paradigme déjà éprouvé, décrit dans la thèse de Paul Blain Lévy [4]. Dans CBPV, toutes les valeurs et fonctions sont stockées dans une pile. Lorsqu'une fonction est appelée, ses arguments sont placés sur la pile avant que la fonction elle-même ne soit ajoutée à la pile. Ce mécanisme permet de suivre de manière explicite les quantités de mémoire pour chaque valeur introduite/éliminée ou fonction appelée/terminée. Aussi, le langage permet d'exprimer clairement les stratégies d'évaluation utilisées dans le code source : soit en *call-by-value*, en évaluant les arguments avant de lancer l'opération, soit en *call-by-name*, en évaluant les arguments uniquement lorsqu'ils seront effectivement utilisés dans la fonction appelée. Ainsi, on fixe quand les évaluations se déroulent, afin de mieux prédire la consommation de mémoire à chaque étape du programme. Ces traits font de CBPV un langage de choix à analyser pour Autobill.

L'entrée est donc imposée. Ainsi, pour étendre l'usage d'Autobill à un autre langage de programmation, un travail de traduction de ce langage donné vers CBPV doit avoir lieu. Cela implique donc de comprendre le langage que l'on compile, notamment les stratégies d'évaluations implicites mises en œuvre, et de l'adapter aux caractéristiques uniques de CBPV citées plus haut.

À partir d'une entrée en CBPV, Autobill traduit le programme en un code machine avec continuations explicites, exprimant explicitement les contraintes de taille qui s'appliquent sur l'entrée. Il l'internalise, c'est à dire construit l'arbre syntaxique abstrait (AST) de ce programme. Ensuite, Autobill infère dans l'AST le typage de ses expressions ainsi que leurs polarités, pour démarquer les calculs et les valeurs dans l'AST. Enfin, en sortie, on remarque dans la figure 1 la possibilité de tirer une interprétation du programme, mais surtout de récupérer les contraintes dans un format MiniZinc [5] ou Coq [6]. Ce sont des outils des assistants de preuve qui permettent, à l'aide d'un langage dédié, d'exprimer explicitement des contraintes logiques. Autobill s'en sert pour décrire les bornes mémoires nécessaires au fonctionnement d'un programme. On peut alors traiter ces équations avec des solveurs, fournis aussi par ces deux outils, pour prouver des propriétés de complexité temporelle ou spatiale.

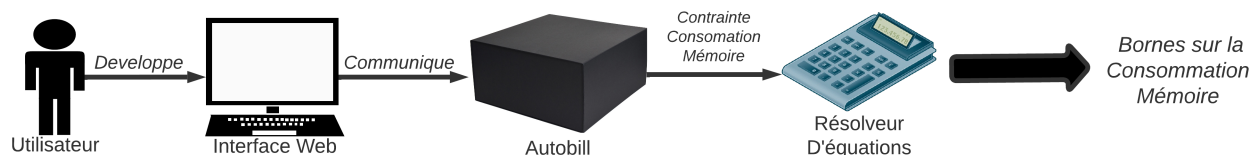


FIGURE 2 – Représentation du système cible

### 1.3 Objectifs du projet

Notre démarche se rapproche de celle de RAML [3] avec leur site officiel : offrir une interface Homme-Machine accessible à tous et illustrant un sujet de recherche en analyse statique.

Le sujet de notre projet STL va donc être de soutenir l’effort de développement en proposant une interface sur le Web permettant la libre manipulation de l’outil Autobill par des utilisateurs à travers un environnement de développement sur navigateur. On souhaite aussi faciliter l’utilisation de l’outil avec un langage fonctionnel pur en entrée plus accessible, un **MiniML**. Cela nous contraint donc à adapter cette nouvelle entrée pour qu’elle soit compatible avec Autobill. Enfin, on se charge aussi de traiter les différentes sorties standards et d’erreurs d’Autobill, notamment les expressions de contraintes, afin de les passer à des solveurs externes, en tirer des preuves de complexité et les afficher directement sur le client Web.

Par rapport à la chaîne d’instructions d’Autobill et à la Figure 1, on se place donc en amont du code LCBPV en entrée et après la sortie en code MiniZinc/Coq.

Notre charge de travail doit se diviser en plusieurs tâches principales :

- L’implémentation du langage MiniML et sa traduction vers CBPV
- La mise en place d’une interface Web
- La mise en relation entre l’interface Web et la machine Autobill
- Le traitement des contraintes d’Autobill par un solveur externe
- Les tests de performances et comparaisons avec les solutions existantes

### 1.4 Processus de conception

Lors de la conception de l’interface, les contraintes étaient multiples. La première était l’interopérabilité des technologies du projet. En effet **Autobill** étant développé en **OCaml**, il était nécessaire de trouver des moyens pour l’adapter à un environnement Web. La seconde était qu’il fallait développer cette interface en simultanément avec **Autobill** et ajuster notre travail en fonction des besoins courants de nos encadrants. Mais la plus importante d’entre elles était le souhait de nos encadrants que l’application soit principalement côté client afin de simplifier son déploiement.

Une fois ces contraintes établies, nous avons dû, tout au long de ce projet, effectuer des choix, que ce soit en matière de design ou de technologies. Nous tenons donc à travers ce rapport à mettre en lumière ces décisions, tout en décrivant le travail qu’elles ont engendré.

## 2 Interface web

Dans l’optique de ne pas se restreindre dans l’utilisation d’outils notamment au niveau du solveur de contraintes, le groupe s’est orienté vers deux structures de projets différentes et indépendantes : l’une fonctionnant avec un client unique, la seconde avec un serveur dédié et un client qui expose ce serveur.

L’avantage réside dans le fait que, lors du développement, si un nouvel outil est amené à être utilisé mais ne dispose de compatibilité sur navigateur Web, alors le serveur peut répondre à ce problème. C’est aussi un sujet de comparaison intéressant à présenter par la suite, que ce soit au niveau des performances que du déploiement de ces solutions.

De cette démarche, il en résulte un code source d’environ 500 lignes, client et serveur compris, faisant tourner notre IDE en ligne dans un état fonctionnel.

### 2.1 Client

Une première approche tout client a été mise œuvre dès le début du projet. Celle-ci permettait de garantir une facilité dans le déploiement en ligne de notre solution. Cette partie se concentra sur la présentation de notre client Web, de l’implémentation des fonctionnalités importantes et des problèmes rencontrés ainsi que leurs résolutions.

#### 2.1.1 Interopérabilité Web-Ocaml

Autobill est un projet entièrement codé en Ocaml, un langage de programmation multiparadigme compilé, et notre projet STL impose un environnement sur navigateur Web, fonctionnant exclusivement avec son langage de script Javascript. Avant de commencer tout codage, il est nécessaire de passer en revue l’état de l’art autour de la compilation et d’exécution de programmes Ocaml sur le Web. De ces recherches vont découler des choix de conception qui vont nous impacter tout le long du semestre.

D’abord, il y’avait la piste des compilateurs vers WebAssembly. WebAssembly est un standard du W3C (World Wide Consortium) qui regroupe un bytecode (.wasm) et un environnement d’exécution compatible avec les navigateurs modernes Javascript. Le bytecode étant une représentation très bas niveau de l’exécutable, l’idée principale est de compiler un langage de programmation plus haut niveau, comme C/C++, Rust,... vers bytecode. Ce bytecode est par la suite compilé par l’environnement d’exécution dans le langage machine de l’hôte. Tant que le support d’un langage de programmation est garanti par un compilateur WebAssembly, il est possible de programmer des applications web complètes dans le langage de son choix ou réinvestir des programmes codés dans ces langages dans une application codée avec des technologies Web.

Pour le cas d’Ocaml, cette possibilité est offerte grâce au post-processeur Wasicaml : il offre un binaire qui prend en entrée un bytecode Ocaml et un fichier .wasm de destination pour traduire chaque instruction dans son équivalent en WebAssembly. Ainsi, on pourrait générer nos fichiers .wasm à partir des binaires d’Autobill et de notre compilateur MiniML, les intégrer à notre application Web et les appeler avec Javascript en leur fournissant les options nécessaires. Cette solution demande néanmoins une dépendance à un projet encore en phase expérimentale, avec certains traits d’Ocaml non intégrés (notamment certaines fonctions du module Unix).

```

1 $ ocamlc -o hello hello.ml
2 $ wasicaml -o hello_wasm hello

```

FIGURE 3 – Chaîne d'instructions pour générer le bytecode WebAssembly

La seconde piste émise par nos tuteurs de projet était l'utilisation de compilateurs Ocaml vers Javascript. La présentation du projet nous a notamment pointé vers `Js_of_Ocaml`. C'est une librairie contenant, entre autres, un compilateur qui transpile du bytecode OCaml en Javascript et propose une grande variété de primitives et de type pour manipuler des éléments Javascript depuis OCaml. L'API de `Js_of_OCaml` est suffisamment fournie pour développer entièrement des applications web complètes et fonctionnelles. On profite alors de l'expérience développeur offerte par Ocaml avec son style de programmation fonctionnelle en y intégrant les traits nécessaires pour construire des pages Web dynamiques.

Il est aussi possible de l'utiliser en coopération avec une base de code Javascript déjà rédigée grâce à la fonctionnalité d'export offerte par `Js_of_Ocaml`. En effet, on pourrait avoir dans un objet Javascript plusieurs méthodes correspondant chacune à un mode d'exécution différent d'Autobill. Chaque méthode prend en entrée le code MiniML à traiter et réalise les transformations nécessaires pour générer la sortie demandée. On peut aussi tirer profit des exceptions d'Ocaml et de la capture des sorties d'erreurs offerte par `Js_of_Ocaml` pour rediriger les messages d'erreurs et les afficher à l'utilisateur. Une fois l'objet exporté et le fichier Ocaml compilé vers Javascript, n'importe quel fichier Javascript du projet pourra ensuite importer l'objet et accéder aux méthodes définies précédemment. On garde ainsi une liberté sur le choix de technologies pour construire notre application.

Enfin, une dernière piste suggérée a été l'utilisation de langages de programmation camélien

```

1 Js.export
2   "ml"
3   (object%js
4     method translate code =
5       let stderr_buff = Buffer.create 100 in
6       Sys_js.set_channel_flusher stderr (Buffer.add_string
7         stderr_buff);
8       let lexbuf = Lexing.from_string ~with_positions:true (Js.
9         to_string code) in
10      let res = string_of_lcbpv_cst (translate_ML_to_LCBPV
11        lexbuf)
12      object%js
13        val resultat = Js.string res
14        val erreur = Js.string (Buffer.contents stderr_buff)
15      end
16    end)

```

FIGURE 4 – Création et exportation d'un objet Javascript avec une méthode de traduction de MiniML vers CBPV

compilables vers Javascript, comme ReasonML ou Rescript. En effet, ce sont tout deux des langages qui ont émergé d'Ocaml et permettent de créer dans un paradigme fonctionnel des applications web complexes. Ils profitent d'une syntaxe ML, d'outils et d'un support communautaire intéressant, faisant des deux langages des alternatives intéressantes pour le Web. Néanmoins, notre objectif principal est la manipulation de la librairie d'Autobill ainsi que celle de MiniML depuis le Web. La compatibilité avec les librairies en Ocaml n'est cependant pas garantie dans le contexte d'une compilation vers Javascript.

Notre groupe a donc opté pour un choix de conception assez flexible afin de nous adapter aux situations que l'on pourrait rencontrer dans le projet. Le client fonctionnera conjointement grâce à une partie Ocaml servie à l'aide de Js\_of\_Ocaml et une partie en Javascript, plus précisément en React.js, pour assurer les fonctionnalités propres à notre IDE.

### 2.1.2 L'éditeur de code

L'idée initiale était de proposer une interface de développement intégrée sur le Web. Celle-ci devait proposer les outils nécessaires pour réaliser les trois tâches suivantes : analyser, écrire et déboguer du code. Pour répondre à ces besoins, plusieurs composants essentiels vont devoir fonctionner ensemble pour fournir l'expérience d'un IDE classique :

- Un éditeur de code pratique et adapté à l'écriture de code dans notre langage MiniML.
- Un menu de navigation entre les programmes proposés par l'application ou créés par l'utilisateur
- Un menu de sélection entre différentes configurations d'analyses
- Une sortie standard et d'erreur pour afficher le retour de l'analyse d'Autobill

La question de l'éditeur de code s'est rapidement posée pour nous. En effet, c'est sur ce composant que va être concentrée l'expérience de l'application et nous avons la volonté de nous rapprocher au plus d'une prise en main similaire aux IDE modernes (Visual Studio Code, Eclipse, IntelliJ...). L'idée d'utiliser un simple champ de texte agrandi n'était donc pas envisageable.

En premier lieu, l'effort de développement a été mis dans la construction *from Scratch* d'un éditeur de code. Cela a permis de fixer clairement les exigences et les fonctionnalités que nous attendions d'un éditeur de code moderne. Déjà, il devait proposer une expérience d'écriture correcte avec toutes les manipulations usuelles et attendues : le comptage des lignes, la recherche et le remplacement rapide d'une suite de caractères (Ctrl+F), le *linting* d'erreurs... Enfin, il y'a le support du langage d'écriture des programmes. Par cela, il est question notamment de coloration syntaxique pour isoler les mots clés et symboles réservés du langage mais aussi de l'autocomplétion basique.

L'examen de nos besoins nous a donc contraints à repousser l'idée d'une création d'un éditeur en partant de zéro. Délivrer les fonctionnalités citées plus haut tout en assurant la stabilité de la solution serait complexe et concentrerait beaucoup d'efforts sur ce composant, au détriment de chantiers clés comme le MiniML. C'est un élément de l'application d'autant plus important qu'il ne doit pas être négligé et demande à lui seul l'effort et le temps nécessaires à un projet universitaire. pour être réalisé. Nous nous sommes donc tournés vers des librairies et solutions déjà existantes afin de les intégrer dans notre projet.



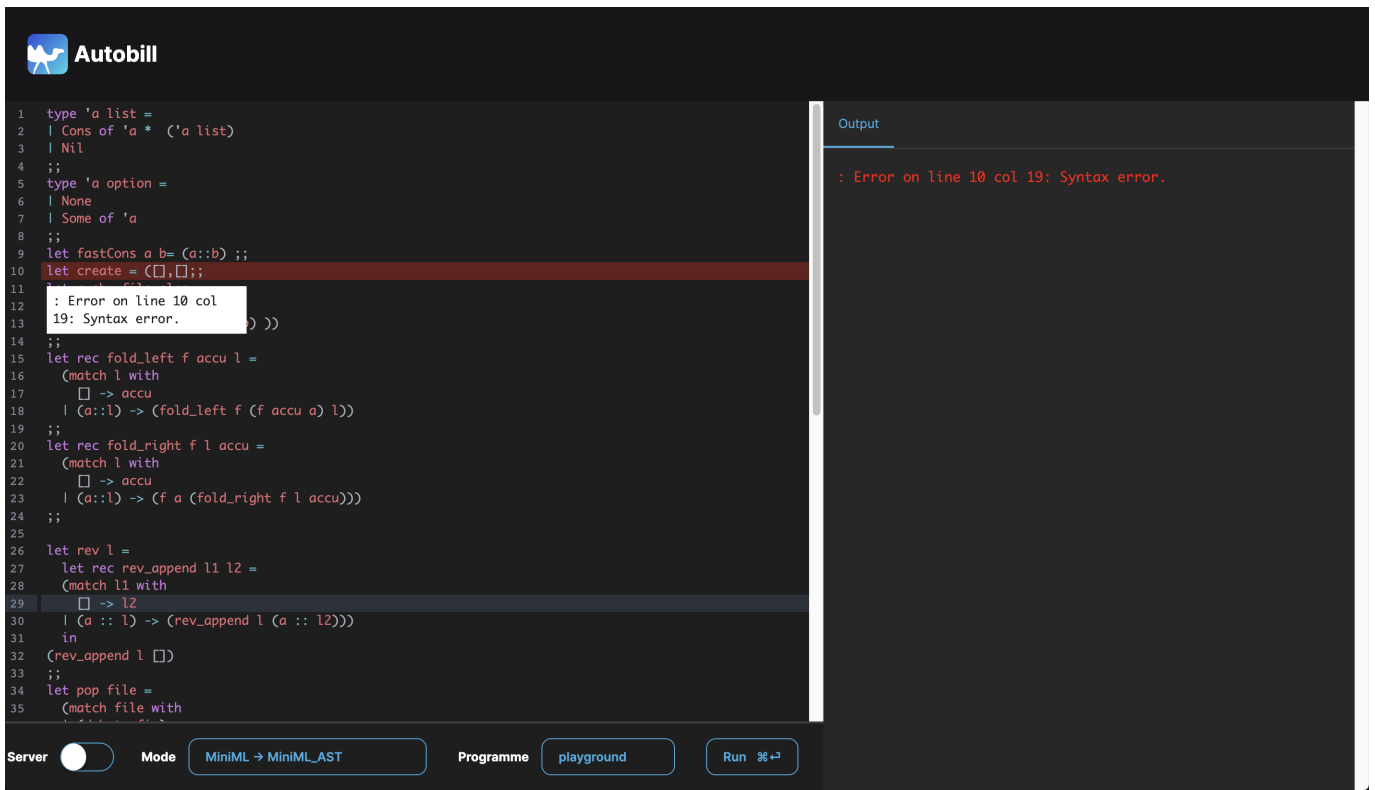


FIGURE 5 – Interface pour Autobill, dans le scénario d’une erreur

Parmi les bibliothèques disponibles, nous avons choisi CodeMirror. C’est un éditeur de code pour navigateur Web, disponible en libre de droits, qui peut s’intégrer à n’importe quelle application Web. Ses fonctionnalités sont exhaustives et il offre de multiples choix de personnalisations pour l’adapter à un langage spécifique. Nativement, il y’a le support intégré pour la syntaxe de langage à la ML comme Ocaml ou F# : MiniML étant un substrat, on peut aisément créer notre fichier de configuration pour ne garder que les mots clés essentiels de notre langage et Codemirror s’occupe du parsing du code et de la coloration de chaque caractère.

Enfin, le *linting* est disponible et permet de dynamiquement surligner des lignes de l’éditeur. Comme illustré dans la Figure 3, on peut, à la volée, afficher directement les parties du programme concernées par une erreur de l’analyse. En effet, quand une méthode de Js\_of\_Ocaml est appelée et qu’une exception est levée durant l’exécution, sa sortie d’erreur peut être redirigée et retournée directement au client. Côté Ocaml, on peut alors formater la sortie d’erreur avec les données de l’exception : la ligne de début et de fin, le type d’erreur rencontrée, la phase de l’analyse durant laquelle l’exception a été levée...

Côté Web, on peut parser la chaîne de caractères et récupérer le JSON parsé pour manipuler les informations sur l’erreur rencontrée (voir Figure 4). CodeMirror offre des fonctions permettant de manipuler l’état global de l’éditeur : son thème, son contenu, le langage de la coloration syntaxique, les options activées... On y retrouve surtout la liste des lignes qui composent l’éditeur. On peut alors leur affecter des changements de style pour les mettre en lumière ou faire apparaître une infobulle avec le message d’erreur si on passe la souris dessus...

```

1 {
2   "loc": {
3     "beginning": { "line": 10, "column": 19},
4     "end": {"line": 10, "column": 21}
5   },
6   "info": "Error on line 10 col 19: Syntax error."
7 }

```

FIGURE 6 – Sortie d’erreur Ocaml parsée

On offre ainsi les outils nécessaires à l’utilisateur pour bien écrire et déboguer son code MiniML. Néanmoins, il faut encore compléter la *pipeline* en traitant toutes les sorties possibles d’Autobill, notamment les équations sur les contraintes mémoires du programme analysée qu’il va falloir résoudre.

### 2.1.3 La résolution de contraintes

Autobill exprime les contraintes mémoires qui s’appliquent à un programme dans des formats proches d’équations mathématiques. Il propose de type de sortie : une sous Coq et une sous MiniZinc. Nous nous concentrerons principalement sur MiniZinc dans cette partie, Coq n’étant pas encore totalement opérationnel et demandant des ressources qui dépassent le cadre du Master d’informatique.

MiniZinc est un langage permettant de décrire des problèmes de manière déclarative à l’aide de contraintes logiques. Il fournit des éléments de syntaxe et de notations très similaires à la programmation et aux mathématiques pour rédiger des modèles qui, lorsqu’ils seront passés à des compilateurs et des solveurs dédiés, vont répondre à des problématiques d’optimisation ou de satisfiabilité. Dans notre cas, on s’intéresse à l’optimisation : on veut déterminer les bornes mémoires minimum du programme passée à Autobill. Néanmoins, on rencontre un problème similaire qu’avec Autobill : la librairie est codée en C++, langage compilé, incompatible avec l’environnement Web. Là encore, l’option de WebAssembly est envisageable avec un compilateur comme Emscripten qui supporte la traduction de code C/C++ vers bytecode WebAssembly. Ici, la librairie core de MiniZinc est beaucoup plus fournie et ne contient pas un point d’entrée direct explicite, comme avec le binaire d’Autobill. La compilation devient alors une tâche plus compliquée et la manipulation sûre de l’outil n’est pas non plus garantie par la suite.

```

1   int: n;
2   var 1..n: x; var 1..n: y;
3   constraint x+y > n;
4   solve satisfy;

```

FIGURE 7 – Exemple de modèle en MiniZinc

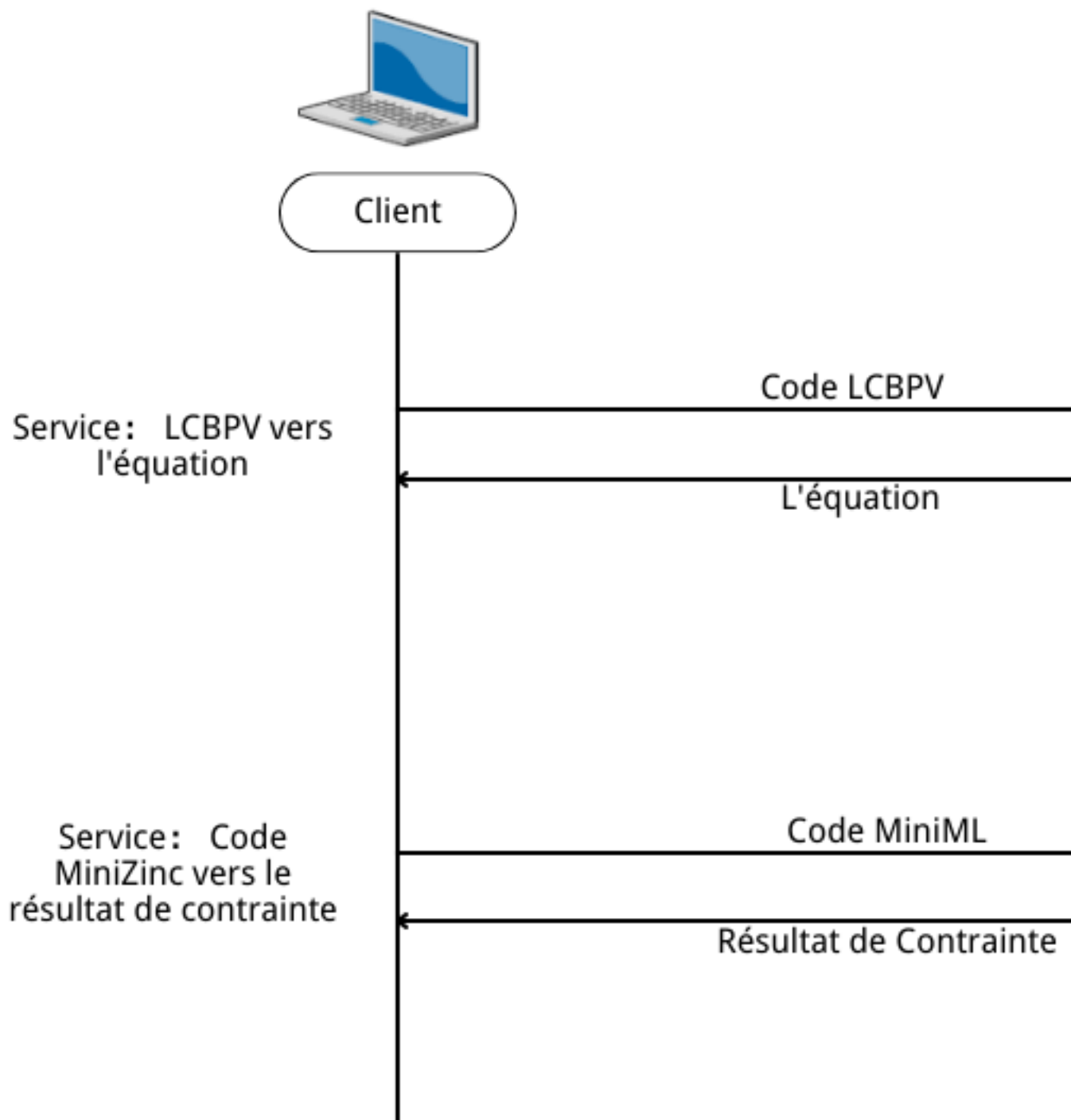
```
1   const model = new Model()
2   model.addFile("playground.mzn", code)
3   const solve = await model.solve({
4     options: {
5       solver: "gecode",
6     },
7   })
```

## 2.2 Serveur + client

Dans le stade actuel d'Autobill, une architecture avec un client seul peut répondre aux exigences du projet STL. Néanmoins, au fur et à mesure que le projet avançait, nous nous sommes rendu compte d'un problème. Autobill évolue constamment et rien ne garantit que ses itérations suivantes puissent être supportées par notre solution. Par exemple, la utilisation de la librairie core de Mini-Zinc peut échouer en mode client unique. Dans l'optique de rendre notre solution plus flexible et *futureproof*, une nouvelle version de notre interface, qui déporte les tâches complexes vers un serveur distant, a été développée.

On a souhaité aussi adapter le client pour qu'il opère dans ces deux architectures différentes. Ainsi, dans notre environnement de développement, on peut facilement faire la bascule entre un mode de fonctionnement local/synchrone et un mode distant/asynchrone.

### 2.2.1 MiniML



11  
FIGURE 8 – Schéma de communication

```

1  const fs = require('fs');
2  const tmpFile = './temp.mzn';
3  fs.writeFileSync(tmpFile, code);
4  const cmd = `minizinc --solver Gecode ${tmpFile}`;
5  exec(cmd, (error, stdout, stderr) => {
6      fs.unlink('temp.mzn', (err) => {
7          if (err) throw err;
8      });
9  })

```

FIGURE 9 – Exécution de code Minizinc

Nous voulons diviser notre architecture Client-Server en deux parties, la partie MiniML et la partie résolution de contraintes. Dans la partie MiniML, nous avons mis en place 5 services principaux en utilisant le protocole HTTP (voir Figure 8) : *’/MiniML/toMiniMLAST’*, *’/MiniML/toLCBPV’*, *’/MiniML/toAutobill’*, *’/MiniML/toAutobillType’* et *’/MiniML/toEquation’*. Pour le service *’/MiniML/toMiniMLAST’*, le client envoie le code du MiniML au serveur via une requête POST. Le serveur le transforme vers l’AST de MiniML et le renvoie au client. Pour le service *’/MiniML/toLCBPV’*, le client envoie le code du MiniML au serveur via une requête POST. Le serveur le transforme vers l’AST de LCBPV et le renvoie au client. Pour le service *’/MiniML/toAutobill’*, le client envoie le code du MiniML au serveur via une requête POST. Le serveur le transforme vers le code d’Autobill et le renvoie au client. Pour le service *’/MiniML/toAutobillType’*, le client envoie le code du MiniML au serveur via une requête POST. Le serveur le transforme vers le code d’Autobill typé et le renvoie au client. Pour le service *’/MiniML/toAutobillType’*, le client envoie le code du MiniML au serveur via une requête POST. Le serveur le transforme vers l’équation résultant de l’analyse statique et le renvoie au client.

## 2.2.2 La résolution de contraintes

Dans la partie résolution de contraintes, nous avons un service *’/minizinc/gecode’* qui exécute le code MiniZinc. Le client envoie le code de MiniZinc au serveur via une requête POST. Le serveur écrit le code dans un fichier temporaire puis l’exécute à l’aide de l’exécutable de ligne de commande `minizinc` avec le solveur Gecode (voir Figure.9), et renvoie le résultat au client.

## 3 MiniML

### 3.1 Description de MiniML

**MiniML** émerge du choix par nos encadrants de créer un langage fonctionnel simple et accessible pour les utilisateurs d'Autobill servant d'abstraction à **CBPV**. Dans le cadre de ce projet MiniML, dispose d'une implémentation écrite en **OCaml**. Nous avons pris la décision de rendre la syntaxe MiniML parfaitement compatible avec OCaml simplifiant les comparaisons avec d'autres outils utilisant **OCaml** en entrée.

Le développement de **MiniML** suivant les besoins de nos encadrants celui-ci est sans effets de bord.

### 3.2 Grammaire

Voici la grammaire BNF du langage MiniML. Pour des raisons de lisibilité, tout les éléments de sucre syntaxique ne seront pas détaillés ici. Les extensions non compatibles avec Ocaml mais compatible avec **Autobill** sont en **Orange**.

#### Elements Nommés

$$\begin{aligned} \langle \text{Id} \rangle &::= [ \text{'a' - z' 'A' - Z' '0' - 9' \_'} ]^+ \\ \langle \text{ConstructeurId} \rangle &::= [ \text{'A' - Z' } \langle \text{Id} \rangle \\ \langle \text{Vartype} \rangle &::= \text{'\"'} [ \text{'a' - z' } [ \text{'0' - 9' } \end{aligned}$$

#### Programmes

$$\begin{aligned} \langle \text{Prog} \rangle &::= | \langle \text{Expr} \rangle \\ &| \langle \text{Def} \rangle \text{';;'} \langle \text{Prog} \rangle \end{aligned}$$

#### Definitions

$$\begin{aligned} \langle \text{Def} \rangle &::= | \text{'let' } \langle \text{Id} \rangle \text{'=' } \langle \text{Expr} \rangle \\ &| \text{'type' } \langle \text{Vartype} \rangle^* \langle \text{Id} \rangle \text{'=' } \langle \text{NewConstructors} \rangle \\ \langle \text{NewConstructors} \rangle &::= | \langle \text{ConstructeurId} \rangle \text{'of' } \langle \text{Type} \rangle \\ &| \langle \text{NewConstructors} \rangle \text{'/' } \langle \text{NewConstructors} \rangle \\ \text{Debut Extension } &| \langle \text{ConstructeurId} \rangle \text{'of' } \langle \text{Type} \rangle \text{'with' } \langle \text{ParamAssigns} \rangle \\ \langle \text{ParamAssigns} \rangle &::= | \langle \text{VarType} \rangle \text{'=' } \langle \text{ParmExpr} \rangle \\ &| \langle \text{ParamAssigns} \rangle \text{' ,' } \langle \text{ParamAssigns} \rangle \end{aligned}$$

<ParmExpr> ::=		<i>1</i>
		<VarType>
		<ParmExpr> <i>+</i> <ParmExpr>
<i>Fin Extension</i>		<ParmExpr> <i>*</i> <ParmExpr>

## Expressions

<Litteral> ::=		[ <i>0 - 9</i> ]+
		[ <i>true</i>   <i>false</i> ]
		<i>( ' ' )</i>

<Expr> ::=		<Litteral>
		<Id>
		<UnaryOperator> <Expr>
		<Expr> <BinaryOperator> <Expr>
		<Expr> <Expr>
		<Expr> <i>,</i> <Expr>
		<i>let</i> <Id> <i>=</i> <Expr> <i>in</i> <Expr>
		<i>fun</i> <Id> <i>-&gt;</i> <Expr>
		<i>fun</i> <i>rec</i> <Id> <Id> <i>-&gt;</i> <Expr>
		<ConstructeurId> <Expr>
		<i>match</i> <Expr> <i>with</i> <MatchCase>

<UnaryOperator> ::=	<i>not</i>
---------------------	------------

<BinaryOperator> ::=	[ <i>and</i>   <i>or</i>   <i>+</i>   <i>-</i>   <i>/</i>   <i>%</i>   <i>*</i>   <i>&lt;</i>   <i>&gt;</i>   <i>=</i> ]
----------------------	--

## Filtrage et Motifs

$$\begin{aligned} \langle \text{MatchCase} \rangle & ::= & | & \langle \text{Pattern} \rangle \text{ '->' } \langle \text{Expr} \rangle \\ & & | & \langle \text{MatchCase} \rangle \text{ '/' } \langle \text{MatchCase} \rangle \\ \\ \langle \text{Pattern} \rangle & ::= & | & \langle \text{Literal} \rangle \\ & & | & \langle \text{Id} \rangle \\ & & | & \langle \text{Pattern} \rangle \text{ ',' } \langle \text{Pattern} \rangle \\ & & | & \langle \text{ConstructeurId} \rangle \langle \text{Pattern} \rangle \end{aligned}$$

## Types

$$\begin{aligned} \langle \text{Type} \rangle & ::= & | & \langle \text{Vartype} \rangle \\ & & | & \langle \text{Id} \rangle \\ & & | & \langle \text{Type} \rangle \langle \text{Type} \rangle \\ & & | & \langle \text{Type} \rangle \text{ '->' } \langle \text{Type} \rangle \\ & & | & \langle \text{Type} \rangle \text{ '*'} \langle \text{Type} \rangle \end{aligned}$$

### 3.3 Call-By-Push-Value

Le paradigme de traitement de langage **Call-By-Push-Value** utilisé par **Autobill** permet à l'aide d'une seule sémantique de traiter deux types de stratégies d'évaluation différentes, **Call By Value** utilisée par **OCaml** et **Call By Name** utilisée par **Haskell** pour mettre en place l'évaluation *Lazy*. Dans CBPV, une distinction a lieu entre les calculs et les valeurs permettant de décider en détail comment ceux-ci sont évalués. Nous permettant, lors de la traduction depuis un autre langage, de choisir le type de stratégie utilisée.

### 3.4 Sémantique de traduction

#### Notation

- $\llbracket \text{let } v = e_1 \text{ in } e_2 \rrbracket_{Expr} \rightarrow \text{let } v = \llbracket e_1 \rrbracket \text{ in } \llbracket e_2 \rrbracket$
- $\llbracket \_ \rrbracket_{Expr} \rightarrow$  est la traduction d'un noeud expr du langage MiniML vers le langage LCBPV
  - À l'intérieur des  $\llbracket \_ \rrbracket$ , on trouve les éléments propre au langage MiniML qui sont traduits vers le langage LCBPV.
  - À l'extérieur des  $\llbracket \_ \rrbracket$ , on trouve les éléments propre au langage LCBPV.
- $X_n$  est le n-ième sous-noeud de l'arbre de syntaxe abstrait.  
Selon les cas  $X$  peut être :
  - $e$  pour les expressions
  - $p$  pour les motifs



- $t$  pour les types
- $d$  pour les définitions
  - $dc$  pour les définitions de constructeurs
- $v$  pour les elements nommées
- $op$  pour les opérateurs
- $l$  pour les litteraux
- $X_1 \dots X_n$  est la liste des sous-noeuds de type  $X$  allant de 1 à  $n$

## Programmes

Un programme MiniML est une suite de taille arbitraire de définitions suivie d'une expression.

$$\llbracket d_1 \dots d_n e \rrbracket_{Prog} \rightarrow \llbracket d_1 \dots d_n \rrbracket return \llbracket e \rrbracket$$

## Définitions

On définit l'opération de traduction  $\llbracket \_ \rrbracket_{Def}$  selon les cas de construction des définitions précisées par la règle de grammaire : **<Def>**

On distingue deux noeuds de types **<Def>** :

- Les définitions de **Variables Globales**,
- Les définitions de **Types**.

$$(VARDEF) \quad \llbracket let v = e \rrbracket_{Def} \rightarrow let v = \llbracket e \rrbracket$$

$$(TYPDEF) \quad \llbracket type v_1 \dots v_N = dc_1 \dots dc_N \rrbracket_{Def} \\ \rightarrow data \llbracket v_1 \rrbracket (\llbracket \dots v_N \rrbracket : +) = \llbracket dc_1 \dots dc_N \rrbracket$$

## Définitions de Constructeurs

On définit l'opération de traduction  $\llbracket \_ \rrbracket_{DefConstructors}$  selon les cas de construction précisées par la règle de grammaire : **<NewConstructors>**

On distingue deux noeuds de types **<NewConstructors>** :

- Les définitions de **Constructeurs Classiques** qui sont compatibles avec Ocaml.
- Les définitions de **Constructeurs Equationnels** qui sont une extension spécifique pour **Autobill** incompatible avec Ocaml et dont la traduction n'est pas encore totalement définie.

$$(MLCONSTRUCTDEF) \quad \llbracket v of t \rrbracket_{DefConstructors} \rightarrow v \llbracket t \rrbracket$$

## Types

On définit l'opération de traduction  $\llbracket \_ \rrbracket_{Type}$  selon les cas de construction des types précisées par la règle de grammaire : **<Type>**

On distingue 4 noeuds de types *Type* :

- Les **Variables de types**.
- Les **Applications de types**.
- Les **Lambda**.
- Les **Tuples**.

$$(TVAR) \quad \llbracket v \rrbracket_{Type} \rightarrow v$$

$$(TAPP) \quad \llbracket t_1 t_2 \rrbracket_{Type} \rightarrow \llbracket t_1 \rrbracket \llbracket t_2 \rrbracket$$

$$(TCLOS) \quad \llbracket t_1 \rightarrow t_2 \rrbracket_{Type} \rightarrow Exp(Fun \llbracket t_1 \rrbracket \rightarrow Thunk \llbracket t_2 \rrbracket)$$

$$(TTUPLE) \quad \llbracket t_1 * t_2 \rrbracket_{Type} \rightarrow \llbracket t_1 \rrbracket \llbracket t_2 \rrbracket$$

## Expressions

On définit l'opération de traduction  $\llbracket \_ \rrbracket_{Expr}$  selon les cas de construction des expressions précisées par la règle de grammaire :  $\langle \text{Expr} \rangle$

On distingue 8 noeuds d'intérêts de types  $\langle \text{Expr} \rangle$  :

- Les **Tuples**.
- Les **Appels de fonctions**.
- Les **Fixation**.
- Les **Lambda**.
- Les **Fonctions Recursive**.
- Les **Constructions**.
- Les **Correspondance de motifs**.
- Les **Appels d'opérateurs**.

**Note :** On ne prend pas en compte les **Litteraux** et les **Variables** dans ce descriptif car leurs traductions sont directes.

$$\begin{aligned} (\text{BINARY}) \quad & \llbracket e_1 \text{ op } e_2 \rrbracket_{Expr} \rightarrow \llbracket e_1 \rrbracket \text{ op } \llbracket e_2 \rrbracket \\ (\text{UNARY}) \quad & \llbracket \text{op } e \rrbracket_{Expr} \rightarrow \text{op } \llbracket e \rrbracket \\ (\text{TUPLE}) \quad & \llbracket e_1, e_2 \rrbracket_{Expr} \rightarrow \text{Tuple}(\llbracket e_1 \rrbracket \llbracket e_2 \rrbracket) \\ (\text{LAMBDA}) \quad & \llbracket \text{fun } v \rightarrow e \rrbracket_{Expr} \rightarrow \text{exp}(\text{get} \mid \text{call}(v) \rightarrow \text{thunk} \llbracket e \rrbracket) \\ (\text{FUN REC}) \quad & \llbracket \text{fun rec } v_1 v_2 \rightarrow e \rrbracket_{Expr} \rightarrow \text{exp}(\text{rec } v_1 \text{ is get} \mid \text{call}(v_2) \rightarrow \text{thunk} \llbracket e \rrbracket) \\ (\text{CALL}) \quad & \llbracket e_1 e_2 \rrbracket_{Expr} \rightarrow \{ \\ & \quad \text{open exp } v_1 = \llbracket e_1 \rrbracket \\ & \quad \text{force thunk}(v_2) = (v_1).\text{call} \llbracket e_2 \rrbracket \\ & \quad \text{return } v_2 \\ & \} \\ (\text{BIND}) \quad & \llbracket \text{let } v = e_1 \text{ in } e_2 \rrbracket_{Expr} \rightarrow \{ \\ & \quad \text{let } v = \llbracket e_1 \rrbracket \\ & \quad \text{return } \llbracket e_2 \rrbracket \\ & \} \\ (\text{CONSTRUCT}) \quad & \llbracket v \text{ } e \rrbracket_{Expr} \rightarrow v \llbracket e \rrbracket \\ (\text{MATCH}) \quad & \llbracket \text{match } e \text{ with } p_1 \dots p_N \rrbracket_{Expr} \rightarrow \text{match } \llbracket e \rrbracket \text{ with } \llbracket p_1 \rrbracket \dots \llbracket p_N \rrbracket \text{ end} \end{aligned}$$

## Motifs et Filtrage

On définit l'opération de traduction  $\llbracket \_ \rrbracket_{Case}$  selon les cas de construction des expressions précisées par la règle de grammaire : **<MatchCase>**

On distingue 4 noeuds d'intérêt de types **<MatchCase>** :

- Les patterns sur **Litteraux**.
- Les patterns sur **Variables**.
- Les patterns sur **Tuple**.
- Les patterns sur **Constructeurs**.

$$\begin{aligned} \text{(PATTERN LIT)} \quad & \llbracket l \rightarrow e \rrbracket_{MatchCase} \rightarrow \llbracket l \rrbracket \rightarrow \llbracket e \rrbracket \\ \text{(PATTERN VAR)} \quad & \llbracket v \rightarrow e \rrbracket_{MatchCase} \rightarrow v \rightarrow \llbracket e \rrbracket \\ \text{(PATTERN TUPLE)} \quad & \llbracket v_1, v_2 \rightarrow e \rrbracket_{MatchCase} \rightarrow Tuple(v_1 \ v_2) \rightarrow \llbracket e \rrbracket \\ \text{(PATTERN CONSTR)} \quad & \llbracket v_1 \ v_2 \rightarrow e \rrbracket_{MatchCase} \rightarrow v_1 \ v_2 \rightarrow \llbracket e \rrbracket \end{aligned}$$

On notera que la traduction des patterns ne prend pas en compte les patterns profonds, c'est à dire les patterns imbriqués dans d'autres patterns. En effet, les patterns profonds ne sont pas encore supportés par le langage cible.

## 4 Conclusion

La réalisation de cette interface a fait intervenir un large panel de sujets en lien avec la formation du Master d’informatique STL et mis à profit les connaissances acquises lors de ce semestre. En premier lieu, le cours d’analyse de programme statique [7] pour toute la partie MiniML et le processus de transformation vers CBPV. Puis, les cours de programmation concurrente répartie, réactive et réticulaire [8], notamment pour la partie réticulaire et l’architecture d’applications Web modernes.

Le projet est à un stade d’avancement satisfaisant. Autobill étant encore en phase expérimentale, celui-ci est alimenté continuellement de nouveautés et corrections que l’on doit intégrer. La suite consistera surtout à consolider les bases établies sur tous les aspects du projet présentés dans ce rapport et les adapter aux changements d’Autobill. Aussi, il serait intéressant à titre de démonstration de comparer notre solution avec celle de Jan Hoffmann et l’interface de RAML [3], mentionnée en section 1.

## Références

- [1] H. Suzanne, “Autobill.” <https://gitlab.lip6.fr/suzanneh/autobill>, 2023.
- [2] J. Hoffmann and S. Jost, “Two decades of automatic amortized resource analysis,” *Mathematical Structures in Computer Science*, vol. 32, pp. 729–759, Mar. 2022.
- [3] J. Hoffmann, “Resource aware ml.” <https://www.raml.co/>.
- [4] P. B. Levy, “Thèse sur call-by-push-value.” <https://www.cs.bham.ac.uk/~pbl/papers/thesisqmwphd.pdf>, 2001.
- [5] “Minizinc.” <https://www.minizinc.org/>, 2023.
- [6] INRIA, “The coq proof assistant.” <https://coq.inria.fr/>, 2023.
- [7] P. MANOURY, “Cours aps.” <https://www-apr.lip6.fr/~manoury/Enseignement/2021-22/APS/index.html>, 2021.
- [8] E. Chailloux, “Cours pc2r.” <https://www-apr.lip6.fr/~chaillou/Public/enseignement/2018-2019/pc2r/public/cours/Cours08.pdf>, 2018.