Catala: un langage de programmation pour formaliser la loi socio-fiscale

Denis Merigoux

Institut National de Recherche en Informatique et en Automatique

11 février 2021



Loi et algorithmes

Code Général des Impôts, Article 197, I, 3, b, 3° (circa 2019)

Le taux de la réduction prévue au premier alinéa du présent b est de 20 %. Toutefois, pour [...], le taux de la réduction d'impôt est égal à 20 % multiplié par le rapport entre :

— au numérateur, la différence entre 20 500 €, pour [...], ou 41 000 €, pour [...], et le montant des revenus mentionnés au troisième alinéa du présent b, et; — au dénominateur, 2 000 €, pour [...], ou 4 000 €, pour [...].

Loi et algorithmes

Code Général des Impôts, Article 197, I, 3, b, 3° (circa 2019)

Le taux de la réduction prévue au premier alinéa du présent b est de 20 %. Toutefois, pour [...], le taux de la réduction d'impôt est égal à 20 % multiplié par le rapport entre :

— au numérateur, la différence entre 20 500 €, pour [...], ou 41 000 €, pour [...], et le montant des revenus mentionnés au troisième alinéa du présent b, et; — au dénominateur, 2 000 €, pour [...], ou 4 000 €, pour [...].

Qu'est-ce qu'un algorithme?

Loi et algorithmes

Code Général des Impôts, Article 197, I, 3, b, 3° (circa 2019)

Le taux de la réduction prévue au premier alinéa du présent b est de 20 %. Toutefois, pour [...], le taux de la réduction d'impôt est égal à 20 % multiplié par le rapport entre :

— au numérateur, la différence entre 20 500 €, pour [...], ou 41 000 €, pour [...], et le montant des revenus mentionnés au troisième alinéa du présent b, et;
— au dénominateur, 2 000 €, pour [...], ou 4 000 €, pour [...].

Qu'est-ce qu'un algorithme?

- Décision sans intervention humaine
- Absence d'ambiguïté
- Données quantitatives (revenus, nombre d'enfants, etc.)

```
TAUX = si (non [...1...]) alors 20 %
```

« Le taux de la réduction prévue au premier alinéa du présent b est de 20 %. Toutefois, pour [...(1)...],... »

```
TAUX = si (non [...1...]) alors
  20 %
sinon
  20 % * (
« il est égal à 20 % multiplié par le rapport entre :
— au numérateur, ...;
— au dénominateur, ... »
```

```
TAUX = si (non [...1...]) alors
  20 %
sinon
  20 % * (
      (si [...2,3...] alors 20 500 € sinon 41000 €)
      - REVENUS
```

« au numérateur, la différence entre 20 500 \in , pour [...(2)...], ou 41 000 \in , pour [...(3)...], et le montant des revenus mentionnés au troisième alinéa du présent b »

```
TAUX = si (non [...1...]) alors
  20 %
sinon
  20 % * (
      (si [...2,3...] alors 20 500 € sinon 41000 €)
      - REVENUS
    (si [...4,5...] alors 2000 € sinon 4000 €)
« au dénominateur, 2 000 €, pour [...(4)...], ou 4 000 €, pour
[...(5)...]. »
```

Allocation familiales calculées par chaque CAF départementale, donc 100+ implémentations différentes, « ouvert », COBOL/PACBASE, 6,9M+ lignes de code

Allocation familiales calculées par chaque CAF départementale, donc 100+ implémentations différentes, « ouvert », COBOL/PACBASE, 6,9M+ lignes de code

Taxe d'habitation calculé par la DGFiP, ouvert, C, 10k+ lignes de code

```
Allocation familiales calculées par chaque CAF départementale, donc 100+ implémentations différentes, « ouvert », COBOL/PACBASE, 6.9M+ lignes de code
```

Taxe d'habitation calculé par la DGFiP, ouvert, C, 10k+ lignes de code

Retraites cotisations, pensions, calculées par les URSAAF et/ou les caisses de retraites (42 régimes), fermé,?

```
Allocation familiales calculées par chaque CAF départementale, donc 100+ implémentations différentes, « ouvert », COBOL/PACBASE, 6,9M+ lignes de code

Taxe d'habitation calculé par la DGFiP, ouvert, C, 10k+ lignes de code
```

Retraites cotisations, pensions, calculées par les URSAAF et/ou les caisses de retraites (42 régimes), fermé,?

Chômage Allocations chômages distribuées par Pôle Emploi, fermé, ?

```
Allocation familiales calculées par chaque CAF départementale,
             donc 100+ implémentations différentes, « ouvert »,
             COBOL/PACBASE, 6,9M+ lignes de code
Taxe d'habitation calculé par la DGFiP, ouvert, C, 10k+ lignes de
             code
   Retraites cotisations, pensions, calculées par les URSAAF
             et/ou les caisses de retraites (42 régimes), fermé,?
   Chômage Allocations chômages distribuées par Pôle Emploi,
             fermé,?
Impôt sur le revenu calculé par la DGFiP, ouvert, M/M++,
             100k+ lignes de code
```

```
Allocation familiales calculées par chaque CAF départementale,
             donc 100+ implémentations différentes, « ouvert »,
             COBOL/PACBASE, 6,9M+ lignes de code
Taxe d'habitation calculé par la DGFiP, ouvert, C, 10k+ lignes de
             code
   Retraites cotisations, pensions, calculées par les URSAAF
             et/ou les caisses de retraites (42 régimes), fermé,?
   Chômage Allocations chômages distribuées par Pôle Emploi,
             fermé,?
Impôt sur le revenu calculé par la DGFiP, ouvert, M/M++,
             100k+ lignes de code
```

Les simulateurs de l'administration

Ines simulateur micro-économique de l'INSEE pour le système socio-fiscal, « ouvert », SAS, 36k+ lignes de code

Les simulateurs de l'administration

- Ines simulateur micro-économique de l'INSEE pour le système socio-fiscal, « ouvert », SAS, 36k+ lignes de code
- OpenFisca simulateur socio-fiscal maintenu par Etalab, ouvert, Python, 40k+ lignes de code

Les simulateurs de l'administration

- Ines simulateur micro-économique de l'INSEE pour le système socio-fiscal, « ouvert », SAS, 36k+ lignes de code
- OpenFisca simulateur socio-fiscal maintenu par Etalab, ouvert, Python, 40k+ lignes de code
- Mon-entreprise.fr simulateur de cotisations salariales et de fiscalité d'entreprise maintenu par Etalab, ouvert, Publicodes, 5k+ lignes de code (en développement)

Des échecs industriels récurrents

Louvois (source France Inter)

Logiciel de paie de l'armée, règles très compliquées pour les primes (174 primes différentes), externalisé à SopraSteria. Calcul largement effectué manuellement auparavant.

⇒ 7 ans de dysfonctionnements catastrophiques, familles de soldats sans ressources, etc. Toujours pas opérationnel en 2019.

Des échecs industriels récurrents

Louvois (source France Inter)

Logiciel de paie de l'armée, règles très compliquées pour les primes (174 primes différentes), externalisé à SopraSteria. Calcul largement effectué manuellement auparavant.

⇒ 7 ans de dysfonctionnements catastrophiques, familles de soldats sans ressources, etc. Toujours pas opérationnel en 2019.

La CIPAV (source France Inter)

Caisse de retraite des architectes et géomètres, chargée en 2005 des auto-entrepreneurs. Projet informatique externalisé à une entreprise espagnole.

⇒ Impayés, trop-payés, retraités menacés de saisie par huissiers, etc. Toujours 3700 bugs en 2014.

Les algorithmes sont-ils fidèles à la loi?

Validation du code

Meilleur des cas : tests « cas-types » écrits par des juristes

- à mettre à jour à chaque changement législatif
- couvrent toutes les situations possibles?
- qualité des tests importante, difficile à évaluer

Les algorithmes sont-ils fidèles à la loi?

Validation du code

Meilleur des cas : tests « cas-types » écrits par des juristes

- à mettre à jour à chaque changement législatif
- couvrent toutes les situations possibles?
- qualité des tests importante, difficile à évaluer

OpenFisca bug trouvé : DOM hors Mayotte, 1 enfant, modulation des allocations (#1426)

Les algorithmes sont-ils fidèles à la loi?

Validation du code

Meilleur des cas : tests « cas-types » écrits par des juristes

- à mettre à jour à chaque changement législatif
- couvrent toutes les situations possibles?
- qualité des tests importante, difficile à évaluer

OpenFisca bug trouvé : DOM hors Mayotte, 1 enfant, modulation des allocations (#1426)

Comment être sûr que son code est correct?

Réunir loi et code

Idée de base

Puisque le code doit suivre la loi, écrivons les deux côte à côte afin de pouvoir les comparer.

Réunir loi et code

Idée de base

Puisque le code doit suivre la loi, écrivons les deux côte à côte afin de pouvoir les comparer.

Questions à régler :

- Quel langage de programmation?
- Quel processus de validation du code?
- Quel processus de production de la loi et du code?

Une équipe interdisciplinaire

Nom (Institution)	Statut	Compétence
Denis Merigoux (Inria Paris)	Doctorant	Méthodes formelles
Liane Huttner (Paris I)	Doctorante	Droit du numérique
Nicolas Chataing (ENS Paris)	Stagiaire de M2	Méthodes formelles
Sarah Lawsky (Northwestern Pritzker School of Law)	Professeure de droit	Droit fiscal et logique
Jonathan Protzenko (Microsoft Research)	Chargé de recherche	Méthodes formelles

Le taux de la réduction prévue au premier alinéa du présent b est de 20 %. Toutefois, pour les contribuables dont les revenus du foyer fiscal, au sens du 1° du IV de l'article 1417, excèdent 19 176 €, pour la première part de quotient familial des personnes célibataires, veuves ou divorcées, ou 38 352 €, pour les deux premières parts de quotient familial des personnes soumises à une imposition commune. ces seuils étant majorés le cas échéant dans les conditions prévues au même premier alinéa, le taux de la réduction d'impôt est égal à 20 % multiplié par le rapport entre : [...]

CGI, article 197, I, 4, b, 3° (circa 2019)

Le taux de la réduction prévue au premier alinéa du présent b est de 20 %. Toutefois, pour les contribuables dont les revenus du foyer fiscal, au sens du 1° du IV de l'article 1417, excèdent 19 176 €, pour la première part de quotient familial des personnes célibataires, veuves ou divorcées, ou 38 352 €, pour les deux premières parts de quotient familial des personnes soumises à une imposition commune. ces seuils étant majorés le cas échéant dans les conditions prévues au même premier alinéa, le taux de la réduction d'impôt est égal à 20 % multiplié par le rapport entre : [...]

CGI, article 197, I, 4, b, 3° (circa 2019)

[Cas de base : condition]

Le taux de la réduction prévue au premier alinéa du présent b est de 20 %. Toutefois, pour les contribuables dont les revenus du foyer fiscal, au sens du 1° du IV de l'article 1417, excèdent 19 176 €, pour la première part de quotient familial des personnes célibataires, veuves ou divorcées, ou 38 352 €, pour les deux premières parts de quotient familial des personnes soumises à une imposition commune. ces seuils étant majorés le cas échéant dans les conditions prévues au même premier alinéa, le taux de la réduction d'impôt est égal à 20 % multiplié par le rapport entre : [...]

CGI, article 197, I, 4, b, 3° (circa 2019)

[Cas de base : condition] \Rightarrow [Cas de base : conséquence]

Le taux de la réduction prévue au premier alinéa du présent b est de 20 %. Toutefois, pour les contribuables dont les revenus du foyer fiscal, au sens du 1° du IV de l'article 1417, excèdent 19 176 €, pour la première part de quotient familial des personnes célibataires, veuves ou divorcées, ou 38 352 €, pour les deux premières parts de quotient familial des personnes soumises à une imposition commune. ces seuils étant majorés le cas échéant dans les conditions prévues au même premier alinéa, le taux de la réduction d'impôt est égal à 20 % multiplié par le rapport entre : [...]

CGI, article 197, I, 4, b, 3° (circa 2019)

[Cas de base : condition] ⇒ [Cas de base : conséquence] [Exception : condition]

Le taux de la réduction prévue au premier alinéa du présent b est de 20 %. Toutefois, pour les contribuables dont les revenus du foyer fiscal, au sens du 1° du IV de l'article 1417, excèdent 19 176 €, pour la première part de quotient familial des personnes célibataires, veuves ou divorcées, ou 38 352 €, pour les deux premières parts de quotient familial des personnes soumises à une imposition commune. ces seuils étant majorés le cas échéant dans les conditions prévues au même premier alinéa, le taux de la réduction d'impôt est égal à 20 % multiplié par le rapport entre :

CGI, article 197, I, 4, b, 3° (circa 2019)

```
[Cas de base : condition] ⇒ [Cas de base : conséquence]

[Exception : condition] ⇒ [Exception : conséquence]
```

Un langage de programmation adapté à la législation

Article D521-1 du code de la sécurité sociale

- I Pour l'application de l'article L. 521-1 , le montant des allocations familiales et de la majoration pour âge prévue à l'article L. 521-3 est défini selon le barème suivant : 1° Lorsque le ménage ou la personne a disposé d'un montant de ressources inférieur ou égal au plafond défini au I de l'article D. 521-3, les taux servant au calcul des allocations familiales sont fixés, en pourcentage de la base mensuelle prévue à l'article L. 551-1, à :
- a) 32 % pour le deuxième enfant à charge;

Un langage de programmation adapté à la législation

Article D521-1 du code de la sécurité sociale

I - Pour l'application de l'article L. 521-1 , le montant des allocations familiales et de la majoration pour âge prévue à l'article L. 521-3 est défini selon le barème suivant : 1° Lorsque le ménage ou la personne a disposé d'un montant de ressources inférieur ou égal au plafond défini au I de l'article D. 521-3, les taux servant au calcul des allocations familiales sont fixés, en pourcentage de la base mensuelle prévue à l'article L. 551-1, à :

a) 32 % pour le deuxième enfant à charge;

Le langage Catala

■ Lisible par des juristes ⇒ mots-clés en français champ d'application, définition, etc.

Le langage Catala

- Lisible par des juristes ⇒ mots-clés en français champ d'application, définition, etc.
- Programmation littéraire [1] comparaison facile du code et du texte de loi

Le langage Catala

- Lisible par des juristes ⇒ mots-clés en français champ d'application, définition, etc.
- Programmation littéraire [1] comparaison facile du code et du texte de loi
- Adapté au style de rédaction des textes juridiques sous condition (logique par défaut [3, 2]), inclusion des champs d'application

Le langage Catala

- Lisible par des juristes ⇒ mots-clés en français champ d'application, définition, etc.
- Programmation littéraire [1] comparaison facile du code et du texte de loi
- Adapté au style de rédaction des textes juridiques sous condition (logique par défaut [3, 2]), inclusion des champs d'application
- Langage à domaine spécifique décidabilité, calculs à virgule fixe pour les montants

Cycle en V X

Cycle en V \times

Méthode agile 🗸

Cycle en V X

Méthode agile 🗸

Programmation en binôme

« La programmation en binôme est une méthode de travail dans laquelle deux développeurs travaillent ensemble sur un même poste de travail. La personne qui rédige le code est appelée conducteur. La seconde personne, appelée observateur, assiste le conducteur en décelant les imperfections, en vérifiant que le code implémente correctement le design et en suggérant des alternatives de développement. »

Cycle en V X

Méthode agile 🗸

Programmation en binôme

« La programmation en binôme est une méthode de travail dans laquelle deux développeurs travaillent ensemble sur un même poste de travail. La personne qui rédige le code est appelée conducteur. La seconde personne, appelée observateur, assiste le conducteur en décelant les imperfections, en vérifiant que le code implémente correctement le design et en suggérant des alternatives de développement. »

Le programmeur est le conducteur, le juriste est l'observateur

Déployer du code Catala en production

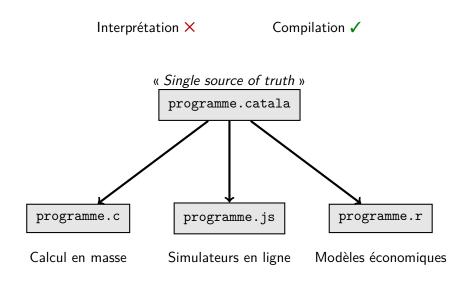
Interprétation X

Déployer du code Catala en production

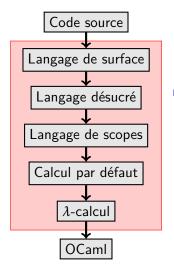
Interprétation X

Compilation \checkmark

Déployer du code Catala en production

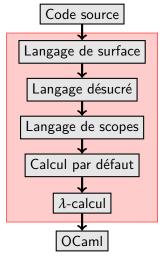


Le compilateur Catala



■ Compilateur en multiples passes

Le compilateur Catala



- Compilateur en multiples passes
- 11 000 lignes de OCaml
- Recherche contributrices et contributeurs!
- github.com/CatalaLang/Catala

Calcul par défaut : syntaxe

$$\langle \varnothing\,,\dots,\varnothing \ | \ \mathtt{false} \, :- \, e \rangle \longrightarrow \varnothing$$

$$\langle \varnothing, ..., \varnothing \mid$$
false $:-e \rangle \longrightarrow \varnothing$

$$\langle \varnothing, \dots, \varnothing \mid \text{true} : - v \rangle \longrightarrow v$$

$$\langle \varnothing, ..., \varnothing \mid \text{false} : -e \rangle \longrightarrow \varnothing$$

$$\langle \varnothing, ..., \varnothing \mid \text{true} : -v \rangle \longrightarrow v$$

$$\langle \varnothing, ..., \varnothing, v, \varnothing, ..., \varnothing \mid e_1 : -e_2 \rangle \longrightarrow v$$

$$\langle \varnothing, ..., \varnothing \mid \text{false} : -e \rangle \longrightarrow \varnothing$$

$$\langle \varnothing, ..., \varnothing \mid \text{true} : -v \rangle \longrightarrow v$$

$$\langle \varnothing, ..., \varnothing, v, \varnothing, ..., \varnothing \mid e_1 : -e_2 \rangle \longrightarrow v$$

$$\frac{\text{nombre_non_vide}(v_1, ..., v_n) > 1}{\langle v_1, ..., v_n \mid e_1 :- e_2 \rangle \longrightarrow \circledast}$$

Formaliser le droit, mais pourquoi?

La loi est-elle cohérente?

Quels seront les effets d'un changement du droit?

Peut-on exploiter des failles de la loi?

Formaliser le droit, mais pourquoi?

La loi est-elle cohérente?

Quels seront les effets d'un changement du droit?

Peut-on exploiter des failles de la loi?

Bac à sable pour les méthodes formelles

- Analyses statiques
- Interprétation abstraite
- Solveurs SMT
- Vérification déductive (Coq, F*, ...)

Modélisation

Caractéristique	Ménage avant	Ménage après	
Revenu annuel	R_0	$R_0 + \Delta_R$	

Modélisation

Caractéristique	Ménage avant	Ménage après	
Revenu annuel Impôt sur le revenu Allocations	R ₀ I ₀ A ₀	$ \begin{vmatrix} R_0 + \Delta_R \\ I_0 + \Delta_I \\ A_0 - \Delta_A \end{vmatrix} $	

Modélisation

Caractéristique	Ménage avant	Ménage après
Revenu annuel	R_0	$R_0 + \Delta_R$
Impôt sur le revenu	<i>I</i> ₀	$I_0 + \Delta_I$
Allocations	A_0	$A_0 - \Delta_A$
Net touché	$N_0 = R_0 - I_0 + A_0$	$N_0 - \Delta_I - \Delta_A$

Modélisation

Caractéristique	Ménage avant	Ménage après
Revenu annuel	R_0	$R_0 + \Delta_R$
Impôt sur le revenu	<i>I</i> ₀	$I_0 + \Delta_I$
Allocations	A_0	$A_0 - \Delta_A$
Net touché	$N_0 = R_0 - I_0 + A_0$	$N_0 - \Delta_I - \Delta_A$

Taux marginal de prélèvement effectif

$$T_{\rm eff} = \frac{\Delta_I + \Delta_A}{\Delta_R}$$

Modélisation

Caractéristique	Ménage avant	Ménage après
Revenu annuel	R_0	$R_0 + \Delta_R$
Impôt sur le revenu	l ₀	$I_0 + \Delta_I$
Allocations	A_0	$A_0 - \Delta_A$
Net touché	$N_0 = R_0 - I_0 + A_0$	$N_0 - \Delta_I - \Delta_A$

Taux marginal de prélèvement effectif

$$T_{\rm eff} = \frac{\Delta_I + \Delta_A}{\Delta_R}$$

Existe-t-il un ménage tel que $T_{\text{eff}} \ge 70\%$?

Contre-exemple (en 2019)



Contre-exemple (en 2019)

Le solveur SMT réfléchit... et trouve!

Couple en concubinage, deux lycéens (15 et 17 ans), à charge deuxième parent (sans activité). Habite en zone II, loyer 897,75 € par mois. Augmentation 250 € par mois.

Contre-exemple (en 2019)

Le solveur SMT réfléchit... et trouve!

Couple en concubinage, deux lycéens (15 et 17 ans), à charge deuxième parent (sans activité). Habite en zone II, loyer 897,75 € par mois. Augmentation 250 € par mois.

Caractéristique	Valeur avant	Valeur après	Variation
Revenu annuel R	33 129,12 €	36 129,12 €	+ 3 000,00 €
IR	3 147,00 €	3 957,00 €	+810,00 €
PA	110,00 €	0,00 €	-110,00 €
AF	132,00 €	132,00 €	0,00 €
ARS	806,00 €	0,00 €	-806,00 €
BC	0,00 €	0,00 €	0,00 €
BL	0,00 €	0,00 €	0,00 €
APL	0,00 €	0,00 €	0,00 €
Net touché N	33 692,12 €	33 756,12 €	+ 64,00 €
Taux marginal			97,9 %

Conclusion

Transformer la loi en code est difficile; les administrations et entreprises utilisent du code de mauvaise qualité

Le langage Catala - https://catala-lang.org

- Programmation littéraire
- Programmation en binôme développeur-juriste
- Valider le code directement en plus des tests

denis.merigoux@inria.fr

Conclusion

Transformer la loi en code est difficile; les administrations et entreprises utilisent du code de mauvaise qualité

Le langage Catala - https://catala-lang.org

- Programmation littéraire
- Programmation en binôme développeur-juriste
- Valider le code directement en plus des tests

denis.merigoux@inria.fr

Démonstration!

References I



Literate Programming.

The Computer Journal, 27(2):97–111, 01 1984.

Sarah B. Lawsky.

A Logic for Statutes.

Florida Tax Review, 2018.

R. Reiter.

Readings in nonmonotonic reasoning.

chapter A Logic for Default Reasoning, pages 68–93. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1987.

Article L521-3 du code de la sécurité sociale

Article L521-3

Chacun des enfants à charge, à l'exception du plus âgé, ouvre droit à partir d'un âge minimum à une majoration des allocations familiales.

```
./securite_sociale_L.catala
268 /*
269 champ d'application CalculAllocationsFamiliales :
270 règle majorations_allocations_familiales.droits_ouverts de enfant
271 sous condition
272 (enfant dans ménage.enfants) et
273 (enfant ≠ ménage.enfant_plus_ågé) et
274 (enfant.åge ≥ 1521_3.åge_limite_alinéa_1 de enfant)
275 conséquence rempli
276 */
```

Toutefois, les personnes ayant un nombre déterminé d'enfants à charge bénéficient de ladite majoration pour chaque enfant à charge à partir de l'âge mentionné au premier alinéa.

```
./securite_sociale_L.catala
279  /*
280  champ d'application CalculAllocationsFamiliales :
281  règle majorations_allocations_familiales.droits_ouverts de enfant
282  sous condition
283  (enfant dans ménage.enfants) et
284  (nombre de ménage.enfants ≥ 1521_3.minimum_alinéa_2) et
285  (enfant.âge ≥ 1521_3.âge_limite_alinéa_1 de enfant)
286  conséquence rempli
287 */
```

Section 132 du code des impôts américains

(c) Qualified employee discount defined

For purposes of this section-

(1) Qualified employee discount

The term "qualified employee discount" means any employee discount with respect to qualified property or services to the extent such discount does not exceed—

(A) in the case of property, the gross profit percentage of the price at which the property is being offered by the employer to customers, or

```
//section_132.catala
32 /*
32 scope QualifiedEmployeeDiscount:
4 definition qualified_employee_discount
5 under condition discount_type with pattern Property consequence
6 equals
7 if employee_discount >
8 customer_price × gross_profit_percentage
9 then customer_price × gross_profit_percentage
4 else employee_discount
4 */
```

(B) in the case of services, 20 percent of the price at which the services are being offered by the employer to customers.