λ-calcul, machines et orthogonalité

Guillaume Munch-Maccagnoni

GdT Logique, cours introductif
24 Octobre 2011*

Le but de cette séance est d'introduire une représentation interactive des preuves et de la réduction en logique intuitionniste. Cela amène naturellement à la technique d'orthogonalité, qui nous donnera un théorème de normalisation ainsi que des propriétés de constructivité pour les preuves.

Il s'agit d'une approche moderne (inspirée de [Kri09, Gir06, CH00]) qui prépare le terrain aux exposés futurs. En effet on la retrouvera pour la logique classique (réalisabilité classique) et pour la logique linéaire (sémantiques dénotationnelles, géométrie de l'interaction).

1. Déduction naturelle et λ-calcul

1.1. Rappel : NJ

Formules

$$A ::= P \mid A \to B \mid A \wedge B \mid A \vee B \mid \top \mid \bot$$

Règles

	introduction	élimination
hypothèse	A	
conjonction	$\dfrac{A \quad B}{A \wedge B}$	$\dfrac{A \wedge B}{A} \quad \dfrac{A \wedge B}{B}$
disjonction	$\dfrac{A}{A \vee B} \quad \dfrac{B}{A \vee B}$	$\dfrac{A \vee B \quad A \to C \quad B \to C}{C}$

*Ajouts du 8 juin 2012 : preuve de la forte normalisation et comparaison avec la méthode de la paire adaptée (Sections A.1 et A.2). Adoption de la terminologie : structure de réalisabilité, $\mathbb{T}, \mathbb{E}$-saturation, par souci de cohérence externe. 9 juillet : corrections mineures.

implication	$\dfrac{\begin{array}{c}[A]\cdots[A]\\ \vdots\\ B\end{array}}{A\to B}$	$\dfrac{A\to B \qquad A}{B}$
vrai	$\dfrac{}{\top}$	
faux		$\dfrac{\bot}{A}$

Réduction

$$\dfrac{\dfrac{\begin{array}{c}[A]\cdots[A]\\ \vdots\\ B\end{array}}{A\to B} \qquad \left.\begin{array}{c}\vdots\\ A\end{array}\right\}\pi}{B} \quad\succ\quad \begin{array}{c}\left.\begin{array}{c}\vdots\\ A\end{array}\right\}\pi \cdots \left.\begin{array}{c}\vdots\\ A\end{array}\right\}\pi\\ \vdots\\ B\end{array}$$

ainsi que d'autres règles pour les autres connecteurs.

1.2. Inteprétation BHK

(Brouwer-Heyting-Kolmogorov) : Une preuve de $A\to B$ est un « algorithme » qui transforme une preuve de A en une preuve de B.

λ-calcul : Une syntaxe des preuves de la déduction naturelle basée sur l'idée qu'une preuve de $A\to B$ est une fonction (au sens des langages de programmation d'ordre supérieur).

- Les hypothèses sont des *variables* $x, y, z\ldots$
- Les arbres de preuves deviennent des *expressions* $t, u, v\ldots$ (arbres syntaxiques).

$$\dfrac{\dfrac{\begin{array}{c}\boldsymbol{x}:[A]\cdots\boldsymbol{x}:[A]\\ \boldsymbol{t(x)}\left\{\begin{array}{c}\vdots\\ B\end{array}\right.\end{array}}{\boldsymbol{x\mapsto t(x)}:A\to B} \qquad \left.\begin{array}{c}\vdots\\ A\end{array}\right\}\boldsymbol{u}}{\boldsymbol{(x\mapsto t(x))(u)}:B} \quad\succ\quad \begin{array}{c}\left.\begin{array}{c}\vdots\\ A\end{array}\right\}\boldsymbol{u} \cdots \left.\begin{array}{c}\vdots\\ A\end{array}\right\}\boldsymbol{u}\\ \boldsymbol{t(u)}:\left\{\begin{array}{c}\vdots\\ B\end{array}\right.\end{array}$$

Grammaire

	introduction	élimination	(analogie)
hypothèse	x		inconnue
conjonction	(t,u)	$\pi_1(t) \quad \pi_2(t)$	$\times$
disjonction	$(1,t) \quad (2,t)$	if t then u else v	$\uplus$
implication	$\lambda x.t$	$t\,u$	$x \mapsto t$
vrai	$()$		

Formellement [1] :

$$t,u ::= x \mid (t,u) \mid (1,t) \mid (2,t) \mid \lambda x.t \mid () \mid \pi_1(t) \mid \pi_2(t) \mid \text{if } t \text{ then } u \text{ else } v \mid t\,u$$

Lien avec la déduction naturelle

Jugements Le jugement de *typage* $t : A$ se lit « t est une preuve de la proposition A » (vocabulaire du logicien) ou encore « t a pour type A » (vocabulaire de l'informaticien).

	introduction	élimination
hypothèse	$x : A$	
conjonction	$\dfrac{t : A \qquad u : B}{(t,u) : A \wedge B}$	$\dfrac{t : A \wedge B}{\pi_1(t) : A} \qquad \dfrac{t : A \wedge B}{\pi_2(t) : B}$
disjonction	$\dfrac{t : A}{(1,t) : A \vee B} \qquad \dfrac{t : B}{(2,t) : A \vee B}$	$\dfrac{t : A \vee B \qquad u : A \to C \qquad v : B \to C}{\text{if } t \text{ then } u \text{ else } v : C}$
implication	$\dfrac{\begin{array}{c} x : [A] \cdots x : [A] \\ \vdots \\ t : B \end{array}}{\lambda x.t : A \to B}$	$\dfrac{t : A \to B \qquad u : A}{t\,u : B}$
vrai	$\dfrac{}{() : \top}$	
faux		$\dfrac{t : \bot}{t : A}$

1. En fait, le if-then-else devra plutôt se lire comme « match t with $(1,x) \to u\,x \mid (2,x) \to v\,x$ » pour les gens qui connaissent CAML.

Réductions

Intuitions :

$$\begin{aligned} \pi_1(t,u) &\succ t \\ \pi_2(t,u) &\succ u \\ \text{if } (1,t) \text{ then } u \text{ else } v &\succ u\,t \\ \text{if } (2,t) \text{ then } u \text{ else } v &\succ v\,t \\ (\lambda x.t)u &\succ t\,[u/x] \end{aligned}$$

Où $t\,[u/x]$ représente le terme t dans lequel chaque occurrence de la variable x est remplacée par u.

Remarque 1. Attention, dans $\lambda x.t$, x est une variable liée, c'est-à-dire que $\lambda x.t$ est le même objet que $\lambda y.t[y/x]$. Par convention, dans $t\,[u/x]$, on supposera toujours que les variables liées de t sont distinctes des variables libres de u.

Malheureusement ce n'est pas une définition complète de la réduction puisqu'il faudrait définir ce que cela signifie de se réduire dans un certain contexte (par exemple pour réduire $\pi_1((\lambda x.t)u)$). Nous ne le faisons pas car cela deviendrait un peu bureaucratique, là où la section suivante introduit une réduction plus simple à manipuler mathématiquement, en rendant explicite la notion de *contexte de réduction*.

1.3. Exemples

Un exemple très simple

Exemple 2. Voici la preuve d'une propriété très simple, $((A \vee B) \to C) \to (A \to C) \wedge (B \to C)$, qui vaut à l'implication son nom catégoriel d'exponentielle (les catégoriciens notent ce morphisme $C^{A+B} \to C^A \times C^B$).

- Preuve en déduction naturelle :

$$\dfrac{\dfrac{\dfrac{[(A \vee B) \to C] \quad \dfrac{[A]}{A \vee B}}{C}}{A \to C} \quad \dfrac{\dfrac{[(A \vee B) \to C] \quad \dfrac{[B]}{A \vee B}}{C}}{B \to C}}{\dfrac{(A \to C) \wedge (B \to C)}{((A \vee B) \to C) \to (A \to C) \wedge (B \to C)}}$$

- Terme du lambda-calcul :

$$\lambda x.(\lambda a.x\,(1,a), \lambda b.x\,(2,b))$$

- Déduction naturelle annotée par des termes du λ-calcul :

$$\dfrac{\dfrac{\dfrac{x : [(A \vee B) \to C] \qquad \dfrac{a : [A]}{(1,a) : A \vee B}}{x\,(1,a) : C}}{\lambda a.x\,(1,a) : A \to C} \qquad \dfrac{\dfrac{x : [(A \vee B) \to C] \qquad \dfrac{b : [B]}{(2,b) : A \vee B}}{x\,(2,b) : C}}{\lambda b.x\,(2,b) : B \to C}}{\dfrac{(\lambda a.x\,(1,a), \lambda b.x\,(2,b)) : (A \to C) \wedge (B \to C)}{\lambda x.(\lambda a.x\,(1,a), \lambda b.x\,(2,b)) : ((A \vee B) \to C) \;\to\; (A \to C) \wedge (B \to C)}}$$

Exercice 3. Si on remplace C par $A \vee B$ ci-dessus on obtient une preuve de :

$$((A \vee B) \to (A \vee B)) \;\to\; (A \to (A \vee B)) \wedge (B \to (A \vee B))$$

de laquelle, par *modus ponens*, on peut en déduire une preuve *indirecte* de $(A \to (A \vee B)) \wedge (B \to (A \vee B))$.

1. Déterminer le terme du λ-calcul correspondant.
2. La réduction termine et donne des termes du λ-calcul correspondant aux preuves *directes* de la conclusion : vérifier cela expérimentalement.

Un autre exemple

Pour l'instant on a parlé de logique propositionnelle et ce n'est pas très expressif. Essayons alors de faire de la théorie des ensembles de façon naïve. On suppose que les formules atomiques (c'est P tout au début, dont je n'ai pas parlé) sont de la forme :

$$a \in b$$

où a et b sont des *individus* de la forme :

- $\alpha, \beta \ldots$ des variables d'individus, ou
- $\{\alpha \mid A(\alpha)\}$ (c'est juste un objet syntaxique) avec $A(\alpha)$ une proposition du langage.

On rajoute à notre système les règles :

$$\dfrac{t : A(a)}{t : (a \in \{\alpha \mid A(\alpha)\})} \qquad \dfrac{t : (a \in \{\alpha \mid A(\alpha)\})}{t : A(a)}$$

Soit alors A une formule quelconque. On considère alors l'individu :

$$E ::= \{\alpha \mid (\alpha \in \alpha) \to A\}$$

Il est notoire que l'on a :

$$\dfrac{E \in E}{(E \in E) \to A} \qquad \dfrac{(E \in E) \to A}{E \in E}$$

L'intérêt de cet exemple réside dans le fait que le paradoxe de Russel, adapté à notre théorie naïve, fournit bel et bien un terme de preuve pour toute formule A :

$$\dfrac{\dfrac{\dfrac{x : [(E \in E) \to A] \qquad \dfrac{x : [(E \in E) \to A]}{x : E \in E}}{x\,x : A}}{\lambda x.x\,x : ((E \in E) \to A) \to A} \qquad \dfrac{\dfrac{x : [E \in E] \qquad \dfrac{x : [E \in E]}{x : (E \in E) \to A}}{x\,x : A}}{\lambda x.x\,x : (E \in E) \to A}}{(\lambda x.x\,x)(\lambda x.x\,x) : A}$$

Cependant, on voit que cette « preuve » ne se réduit qu'en elle-même :

$$(\lambda x.x\,x)(\lambda x.x\,x) \succ (\lambda x.x\,x)(\lambda x.x\,x) \succ \dots$$

Par conséquent, les règles de cette théorie naïve des ensembles ne satisfont pas au critère de normalisation, ce qui est notre argument pour les rejeter.

Slogan (Gentzen, Girard [Gir06]) :

$$\text{incohérence} \Longleftrightarrow \text{existence de preuves qui bouclent}$$

On aura l'occasion de voir à d'autres séances des systèmes cohérents plus expressifs que la logique propositionnelle intuitionniste que j'ai introduite pour l'instant. La technique qu'on présente ici se généralise à des systèmes très puissants ; mais pour cela on a besoin d'une meilleure représentation de la réduction.

2. Machines et calcul des séquents

2.1. Réduction sous contexte

Si on souhaite réduire « if $(\lambda x.t)u$ then v else w » il est nécessaire de d'abord réduire $(\lambda x.t)u$: les réductions doivent donc s'effectuer dans des *contextes de réduction*, qui représentent le lieu où la réduction se poursuit lorsqu'on a fini de réduire ce qu'il y a à l'intérieur. Ici, le contexte est « if $\dots$ then v else w » et le problème est que l'on ne sait pas réduire cela tant que « $\dots$ » ne contient pas $(1, t')$ ou $(2, t')$ avec t' un terme quelconque.

Désormais, on prendra la notion de contexte très au sérieux :

- On notera E les contextes que l'on définit plus bas.
- On ne réduira un terme qu'en présence d'un contexte. La relation de réduction $\succ$ est donc désormais définie uniquement sur des « machines » qui sont des couples $\langle t \mid E\rangle$ constitués d'un terme t et d'un contexte E, que l'on note aussi c :

$$c ::= \langle t \mid E\rangle$$

Réductions adjointes Ce sont les règles d'élimination de la déduction naturelle qui créent les contextes de réduction :

$$\begin{aligned}
\langle \pi_1(u) \,|\, E\rangle &\succ \langle u \,|\, \pi_1 \cdot E\rangle \\
\langle \pi_2(u) \,|\, E\rangle &\succ \langle u \,|\, \pi_2 \cdot E\rangle \\
\langle \text{if } t \text{ then } u \text{ else } v \,|\, E\rangle &\succ \langle t \,|\, (u|v) \cdot E\rangle \\
\langle t\,u \mid E\rangle &\succ \langle t \mid u \cdot E\rangle
\end{aligned}$$

Remarque : les règles d'élimination sur les termes sont toutes de la forme $\langle \tau^*(t) \,|\, E \rangle \succ \langle t \,|\, \tau(E) \rangle$: ce sont donc les « adjointes » de règles d'introduction sur les contextes.

Il faut aussi un contexte « vide » initial, que je note $\star$.

$$E ::= \star \mid \pi_1 \cdot E \mid \pi_2 \cdot E \mid (u|v) \cdot E \mid u \cdot E$$

Réductions principales

$$\begin{aligned}
\langle \lambda x.t \,|\, u \cdot E \rangle &\succ \langle t\,[u/x] \,|\, E \rangle \\
\langle (t,u) \,|\, \pi_1 \cdot E \rangle &\succ \langle t \,|\, E \rangle \\
\langle (t,u) \,|\, \pi_2 \cdot E \rangle &\succ \langle u \,|\, E \rangle \\
\langle (1,t) \,|\, (u|v) \cdot E \rangle &\succ \langle u \,|\, t \cdot E \rangle \\
\langle (2,t) \,|\, (u|v) \cdot E \rangle &\succ \langle v \,|\, t \cdot E \rangle
\end{aligned}$$

Des règles d'introduction sur les contextes plutôt que des règles d'élimination ; des réductions principales entre règles d'introduction de chaque côté ? Cela rappelle le calcul des séquents, pardi !

2.2. Calcul des séquents intuitionnistes, rappels

Un séquent (intuitionniste) est une liste de formules de la forme :

$$A_1, \ldots, A_n \vdash B$$

qui correspond à la proposition :

$$(A_1 \wedge \cdots \wedge A_n) \to B$$

Les A_i sont donc des hypothèses et B est la conclusion. (L'ordre des hypothèses n'a pas d'importance.) On note souvent $\Gamma = A_1, \ldots, A_n$ une liste quelconque d'hypothèses.

	introduction droite	introduction gauche
hypothèse	$\dfrac{}{\Gamma, A \vdash A}$	
coupure	$\dfrac{\Gamma \vdash A \qquad \Gamma, A \vdash B}{\Gamma \vdash B}$	
conjonction	$\dfrac{\Gamma \vdash A \qquad \Gamma \vdash B}{\Gamma \vdash A \wedge B}$	$\dfrac{\Gamma, A \vdash C}{\Gamma, A \wedge B \vdash C} \qquad \dfrac{\Gamma, B \vdash C}{\Gamma, A \wedge B \vdash C}$
disjonction	$\dfrac{\Gamma \vdash A}{\Gamma \vdash A \vee B} \qquad \dfrac{\Gamma \vdash B}{\Gamma \vdash A \vee B}$	$\dfrac{\Gamma, A \vdash C \qquad \Gamma, B \vdash C}{\Gamma, A \vee B \vdash C}$
implication	$\dfrac{\Gamma, A \vdash B}{\Gamma \vdash A \to B}$	$\dfrac{\Gamma \vdash A \qquad \Gamma, B \vdash C}{\Gamma, A \to B \vdash C}$
vrai	$\dfrac{}{\Gamma \vdash \top}$	
faux		$\dfrac{}{\Gamma, \bot \vdash A}$

Exemple 4. Élimination de l'implication :

$$\dfrac{\Gamma \vdash A \to B \qquad \dfrac{\Gamma \vdash A \qquad \dfrac{}{\Gamma, B \vdash B}\,\text{hyp.}}{\Gamma, A \to B \vdash B}\,{\to}\vdash}{\Gamma \vdash B}\,\text{coupure}$$

2.3. Typage des machines

Idée :

$$\dfrac{\Gamma \vdash A \qquad \Gamma, A \vdash B}{\Gamma \vdash B}$$

devient :

$$\dfrac{\Gamma \vdash t : A \qquad \Gamma \mid E : A \vdash \star : B}{\langle t \mid E \rangle : (\Gamma \vdash \star : B)}$$

où Γ devient une liste[2] $x_1 : A_1, \ldots, x_n : A_n$ qui contient les variables non liées des termes, contextes et machines considérés.

Jugements On a désormais trois jugements. Le premier concerne les termes et est similaire à celui de la déduction naturelle :

$$\Gamma \vdash t : A$$

Il se lit « t est une preuve de A sous les hypothèses Γ ». Toutes les hypothèses de la preuve, c'est-à-dire toutes les variables non liées de t doivent apparaître dans Γ.

Le second est le suivant :

$$\Gamma \mid E : A \vdash \star : B$$

Il se lit « le contexte E est une contre-preuve (une réfutation) de A sous les hypothèses Γ et sous la conclusion B ». Le séquent sous-jacent est $\Gamma, A \vdash B$, mais la séparation verticale entre Γ et A est ajoutée pour la lisibilité. En effet, $E : A$ type un contexte tandis que Γ type les variables non liées de E.

Le troisième jugement concerne les machines :

$$\langle t \mid E \rangle : (\Gamma \vdash \star : B)$$

$\langle t \mid E \rangle$ ne possède pas de type en lui-même, mais ses variables (dans Γ) ainsi que sa conclusion en ont.

2. En fait, une fonction partielle des variables dans les formules.

Correspondance machines-calcul des séquents [3]

	introduction droite	introduction gauche
hypothèse	$\dfrac{}{\Gamma, x : A \vdash x : A}$	$\dfrac{}{\Gamma \mid \star : A \vdash \star : A}$
coupure	$\dfrac{\Gamma \vdash t : A \qquad \Gamma \mid E : A \vdash \star : B}{\langle t \mid E \rangle : (\Gamma \vdash \star : B)}$	
adjoint	$\dfrac{c : (\Gamma \vdash \star : B)}{\Gamma \vdash c^* : B}$	
conjonction	$\dfrac{\Gamma \vdash t : A \qquad \Gamma \vdash u : B}{\Gamma \vdash (t,u) : A \wedge B}$	$\dfrac{\Gamma \mid E : A \vdash C}{\Gamma \mid \pi_1 \cdot E : A \wedge B \vdash C} \qquad \dfrac{\Gamma \mid E : B \vdash C}{\Gamma \mid \pi_2 \cdot E : A \wedge B \vdash C}$
disjonction	$\dfrac{\Gamma \vdash t : A}{\Gamma \vdash (1,t) : A \vee B} \qquad \dfrac{\Gamma \vdash t : B}{\Gamma \vdash (2,t) : A \vee B}$	$\dfrac{\Gamma \mid E : C \vdash \star : D \qquad \Gamma \vdash t : A \to C \qquad \Gamma \vdash u : B \to C}{\Gamma \mid (t \mid u) \cdot E : A \vee B \vdash \star : D}$
implication	$\dfrac{\Gamma, x : A \vdash t : B}{\Gamma \vdash \lambda x.t : A \to B}$	$\dfrac{\Gamma \vdash t : A \qquad \Gamma \mid E : B \vdash \star : C}{\Gamma \mid t \cdot E : A \to B \vdash \star : C}$
vrai	$\dfrac{}{\Gamma \vdash () : \top}$	
faux		$\dfrac{}{\Gamma \mid \star : \bot \vdash \star : A}$

Adjoint On remplace dans la grammaire des termes les règles d'élimination par l'adjoint c^* :

$$t, u ::= x \mid (t,u) \mid (1,t) \mid (2,t) \mid \lambda x.t \mid () \mid c^*$$

L'adjoint possède la règle de réduction :

$$\langle c^* \mid E \rangle \succ c\,[E/\star]$$

où la substitution $c\,[E/\star]$ substitue E à l'unique occurrence de $\star$ dans c qui n'est pas sous un adjoint. (On considère donc que les occurrences de $\star$ apparaissant dans c à l'intérieur d'un adjoint c'^* sont liées.)

3. Par simplicité de la machine considérée dans cette note, j'ai dû choisir une règle d'introduction gauche de la disjonction qui diffère de la règle du calcul des séquents. Obtenir une meilleure correspondance entre la machine et le calcul des séquents (et une règle assez petite pour ne pas dépasser du tableau) est possible et va de pair avec un meilleur traitement des connecteurs positifs en logique ; pour cela il est nécessaire d'introduire des « contextes adjoints » (voir [CH00, Mun09], ou le prochain exposé GdT).

On peut alors *définir* les règles d'élimination :

$$\pi_1(t) \stackrel{\text{def}}{=} \langle t \mid \pi_1 \cdot \star \rangle^*$$
$$\pi_2(t) \stackrel{\text{def}}{=} \langle t \mid \pi_2 \cdot \star \rangle^*$$
$$\text{if } t \text{ then } u \text{ else } v \stackrel{\text{def}}{=} \langle t \mid (u|v) \cdot \star \rangle^*$$
$$t\,u \stackrel{\text{def}}{=} \langle t \mid u \cdot \star \rangle^*$$

En effet, avec ces définitions, on obtient non seulement les règles qui vont bien [4], comme :

$$\left\langle \langle t \mid u \cdot \star \rangle^* \,\middle|\, E \right\rangle \succ \langle t \mid u \cdot E \rangle$$

mais ces définitions se déduisent des dérivations canoniques de l'élimination à partir de l'introduction droite, dont celle pour l'implication est donnée plus haut en exemple :

Exemple 5. Élimination de l'implication :

$$\dfrac{\dfrac{\Gamma \vdash t : A \to B \qquad \dfrac{\dfrac{\Gamma \vdash u : A \qquad \overline{\Gamma \mid \star : B \vdash \star : B}^{\text{hyp.}\vdash}}{\Gamma \mid u \cdot \star : A \to B \vdash \star : B}^{\to\vdash}}{\langle t \mid u \cdot \star \rangle : (\Gamma \vdash \star : B)}^{\text{coupure}}}{\Gamma \vdash \langle t \mid u \cdot \star \rangle^* : B}^{\text{adjoint}}$$

Slogan

Réduction des machines $\Longleftrightarrow$ Élimination des coupures

3. Orthogonalité et normalisation

On va associer à chaque formule A un ensemble $|A|$ de termes qui « se comportent », au sens de la réduction, similairement aux preuves de A, ainsi qu'un ensemble $||A||$ de contextes qui « se comportent » similairement aux contextes correspondant.

Les deux ensembles $|A|$ et $||A||$ sont reliés par une notion d'« orthogonalité » : $|A| = ||A||^{\perp}$ et $||A|| = |A|^{\perp}$ – on dira qu'ils sont *polaires* l'un de l'autre. Un terme t et un contexte E sont polaires lorsque leur interaction au sein de la machine $\langle t \mid E \rangle$ produit un calcul « acceptant », ce que l'on définira et que l'on note $t \perp\!\!\!\perp E$. Par conséquent, $|A|$ est l'ensemble des termes qui interagissent « correctement » avec chacun des contextes de $||A||$, et vice-versa.

4. L'élimination de $\bot$, en revanche, s'obtient avec la règle dérivée :

$$\frac{\Gamma \vdash t : \bot}{\Gamma \vdash \langle t \mid \star \rangle^* : A} \text{ au lieu de } \frac{\Gamma \vdash t : \bot}{\Gamma \vdash t : A}$$

mais la différence importe peu, puisque le comportement calculatoire est le même :

$$\left\langle \langle t \mid \star \rangle^* \,\middle|\, E \right\rangle \succ \langle t \mid E \rangle .$$

Exemple 6. Si on voulait formaliser l'arithmétique, on aurait (entre autres) un prédicat $x \in \mathbb{N}$, et une preuve de ce prédicat calculerait, à travers la réduction, un entier correspondant. Par exemple, une preuve de $\exists x \in \mathbb{N}, f(x) = 0$ calculerait un entier n tel que $f(n) = 0$. Dans ce but, on serait conduit à ajouter au langage des termes des constantes $1, 2, \dots$ représentant les entiers. Alors, l'ensemble $\{1, 2, \dots\}^\perp$ contiendrait les contextes qui attendent un entier en entrée. L'ensemble $\{1, 2, \dots\}^{\perp\perp}$ des termes qui interagissent correctement avec ces contextes représenterait par conséquent les programmes qui calculent un entier. C'est donc le bon candidat pour interpréter le type des entiers, ou, en logique, le prédicat $\exists x \in \mathbb{N}$. (Une telle formalisation de l'arithmétique va cependant au-delà de notre programme pour cette séance.)

Contrairement à l'ensemble des preuves de A, l'ensemble $|A|$ est défini par récurrence sur la formule. Néanmoins, le théorème principal que l'on donne plus bas relie la déduction des systèmes logiques à cette interprétation des formules :

Théorème 7 (Lemme d'adéquation, cas particulier)**.**
- *Si* $\vdash t : A$ *alors* $t \in |A|$,
- *Si* $\mid E : A \vdash \star : B$ *et* $\star \in ||B||$, *alors* $E \in ||A||$,
- *Si* $c : (\vdash \star : B)$ *et* $\star \in ||B||$, *alors* $c \in \perp\!\!\!\perp$,

En effet, on peut percevoir déjà que le fait que $|A|$ et $||A||$ sont polaires corrobore la règle de la coupure : si $t \in |A|$ et $E \in ||A||$ alors $\langle t \mid E \rangle \in \perp\!\!\!\perp$.

Avant que cela ait un sens, nous définissons la relation $t \perp\!\!\!\perp E$, ainsi que pour toute formule A, deux ensembles $|A|$ et $||A||$ respectivement de termes et de contextes.

3.1. Pôle

Soit $\mathbb{T}$ un ensemble de termes et $\mathbb{E}$ un ensemble de contextes.

Définition 8 (Ensemble polaire)**.** On suppose donné un *pôle* $\perp\!\!\!\perp$, qui est une partie de $\mathbb{T} \times \mathbb{E}$.
- On note $t \perp\!\!\!\perp E \iff \langle t \mid E \rangle \in \perp\!\!\!\perp$.
- Si $\mathcal{T}$ est un ensemble de termes, on pose $\mathcal{T}^\perp = \{E \in \mathbb{E} \mid \forall t \in \mathcal{T}\ t \perp\!\!\!\perp E\}$.
- Si $\mathcal{E}$ est un ensemble de contextes de réduction, on pose $\mathcal{E}^\perp = \{t \in \mathbb{T} \mid \forall E \in \mathcal{E}\ t \perp\!\!\!\perp E\}$. [5]

Proposition 9 (Propriétés de base des polaires)**.** *Soient* $\mathcal{U}, \mathcal{V}$ *deux sous-ensembles de* $\mathbb{T}$ *ou de* $\mathbb{E}$.

1. *Si* $\mathcal{U} \subseteq \mathcal{V}$ *alors* $\mathcal{V}^\perp \subseteq \mathcal{U}^\perp$,
2. $\mathcal{U} \subseteq \mathcal{U}^{\perp\perp}$,
3. $\mathcal{U}^\perp = \mathcal{U}^{\perp\perp\perp}$,
4. $\left(\bigcup_i \mathcal{U}_i\right)^\perp = \bigcap_i \mathcal{U}_i{}^\perp$.

Démonstration. Exercice. □

5. L'ambiguïté sur le sens de $\emptyset^\perp$, qui vaut ou bien $\mathbb{T}$ ou bien $\mathbb{E}$, sera levée par le contexte.

Remarque 10. Il s'agit en réalité d'une structure très générale dont l'orthogonalité de l'algèbre linéaire n'est qu'un cas particulier : toute relation entre deux ensembles U et V induit une correspondance de Galois antitone entre $\mathcal{P}(U)$ et $\mathcal{P}(V)$ ordonnés par l'inclusion, qui vérifie en outre (4.).

N'importe quels ensembles $\mathbb{T}$ et $\mathbb{E}$ et n'importe quel pôle $\perp\!\!\!\perp$ ne vont pas nécessairement donner une interprétation des formules compatible avec la logique ! On doit imposer des conditions très générales sur le choix de $\mathbb{T}, \mathbb{E}$ et $\perp\!\!\!\perp$ — $\mathbb{T}$ et $\mathbb{E}$ doivent être stables par certaines constructions de base, et le pôle doit être clos par expansion logique, c'est-à-dire l'inverse de la réduction.

Définition 11. Une *structure de réalisabilité* [6] est un triplet $(\mathbb{T}, \mathbb{E}, \perp\!\!\!\perp)$ tel que :

$\mathbb{T}$ un ensemble de termes contenant $()$ et stable par paire et par union disjointe, autrement dit :

1. $() \in \mathbb{T}$,
2. si $t \in \mathbb{T}$ et $u \in \mathbb{T}$ alors $(t, u) \in \mathbb{T}$,
3. si $t \in \mathbb{T}$ alors $(1, t) \in \mathbb{T}$ et $(2, t) \in \mathbb{T}$,

$\mathbb{E}$ un ensemble de contextes contenant $\star$ et stable par empilement d'un terme, autrement dit :

1. $\star \in \mathbb{E}$
2. si $t \in \mathbb{T}$ et $E \in \mathbb{E}$ alors $t \cdot E \in \mathbb{E}$.

Enfin, le pôle $\perp\!\!\!\perp \subseteq \mathbb{T} \times \mathbb{E}$ est $\mathbb{T}, \mathbb{E}$-*saturé*, c'est-à-dire :

1. si $E \in \mathbb{E}$ et $c\,[E/\star] \in \perp\!\!\!\perp$ alors $\langle c^* \,|\, E\rangle \in \perp\!\!\!\perp$,
2. si $t_1, t_2 \in \mathbb{T}$ et $\langle t_i \,|\, E\rangle \in \perp\!\!\!\perp$ alors $\langle (t_1, t_2) \,|\, \pi_i \cdot E\rangle \in \perp\!\!\!\perp$,
3. si $t_1, t_2 \in \mathbb{T}$ et $\langle t_i \,|\, u \cdot E\rangle \in \perp\!\!\!\perp$ alors $\langle (i, u) \,|\, (t_1 | t_2) \cdot E\rangle \in \perp\!\!\!\perp$,
4. si $u \in \mathbb{T}$ et $\langle t\,[u/x] \,|\, E\rangle \in \perp\!\!\!\perp$ alors $\langle \lambda x.t \,|\, u \cdot E\rangle \in \perp\!\!\!\perp$.

Remarquer que la condition de $\mathbb{T}$,$\mathbb{E}$-saturation impose des contraintes sur $\mathbb{T}$ et $\mathbb{E}$, puisqu'il faut s'assurer que $\perp\!\!\!\perp$ est bien une partie de $\mathbb{T} \times \mathbb{E}$.

Dans la suite, on suppose que $(\mathbb{T}, \mathbb{E}, \perp\!\!\!\perp)$ est une structure de réalisabilité.

Exemple 12. On peut prendre pour $\mathbb{T}$ et $\mathbb{E}$ les ensembles des termes et des contextes clos (dont toutes les variables sont liées). Alors, tout ensemble $\mathcal{C}$ de couples $\langle t \,|\, E\rangle$ clos définit une structure de réalisabilité en posant $\perp\!\!\!\perp = \{c \in \mathbb{T} \times \mathbb{E} \,|\, \exists c' \in \mathcal{C},\ c \succ \cdots \succ c'\}$. Par exemple, en prenant $\perp\!\!\!\perp$ l'ensemble des $\langle t \,|\, E\rangle$ clos qui se réduisent en plusieurs étapes sur $\langle (1, u) \,|\, \star\rangle$ ou sur $\langle (2, u) \,|\, \star\rangle$ pour $u \in \mathbb{T}$, on obtiendra à travers le lemme d'adéquation la *propriété de la disjonction* : une preuve de $A \vee B$, après réduction, indique s'il s'agit d'une preuve de A ou d'une preuve de B. (Voir Corollaire 18.)

6. Terminologie empruntée à Krivine, bien qu'il s'agisse d'une variante adaptée au calcul considéré.

Exemple 13. Pour prouver la normalisation (le fait que la réduction des termes typés se termine), on prendra $\mathbb{T}$ l'ensemble des termes qui normalisent (ce que l'on définira), $\mathbb{E}$ l'ensemble des contextes qui normalisent, et $\perp\!\!\!\perp$ l'ensemble des machines qui normalisent. (Remarquer qu'on peut avoir t et E qui normalisent sans avoir $\langle t \mid E\rangle$ qui normalise.) On montrera qu'il s'agit bien d'une structure de réalisabilité, et on déduira le résultat du lemme d'adéquation.

3.2. Analogie avec les automates finis non déterministes

(Pour les initiés.)

Soit $\mathcal{A} = (\Sigma, Q, \delta, q_I, Q_F)$ un automate fini non déterministe. L'analogie suivante avec les machines et l'interprétation polaire peut aider :

termes $t \in \mathbb{T}$	mots $\omega \in \Sigma^*$
contextes $E \in \mathbb{E}$	états $q \in Q$
réductions $\langle t \mid E\rangle \succ \langle t' \mid E'\rangle$	$\langle a.\omega \mid q\rangle \succ \langle \omega \mid q'\rangle$ pour toute transition $q' \in \delta(q,a)$
pôle $\perp\!\!\!\perp$	états finals ($\perp\!\!\!\perp = \left\{\langle \omega \mid q\rangle \;\middle|\; \exists q_F \in Q_F, \langle \omega \mid q\rangle \succ \cdots \succ \langle \varepsilon \mid q_F\rangle\right\}$)
formules A	expressions rationnelles
interprétations $\lvert A\rvert$	langages rationnels

Proposition 14. *Les langages de la forme $S^\perp$ pour $S \subseteq Q$ sont rationnels, et en particulier $\{q_I\}^\perp$ est le langage reconnu par l'automate.*

Démonstration. On prend la convention que si $n = -1$, alors $a_0 \cdots a_n$ est le mot vide. On raisonne par équivalence : $a_0 \cdots a_n \in \{q_I\}^\perp$ signifie $\langle a_0 \cdots a_n \mid q_I\rangle \in \perp\!\!\!\perp$, c'est-à-dire qu'il existe $q_F \in Q_F$ tel que $\langle a_0 \cdots a_n \mid q_I\rangle \succ^* \langle \varepsilon \mid q_F\rangle$, autrement dit qu'il existe $n+2$ états $q_0, \ldots, q_{n+1}$ tels que $q_I = q_0 \xrightarrow{a_0} q_1 \xrightarrow{a_1} \cdots \xrightarrow{a_n} q_{n+1} = q_F$ (dans la notation usuelle des transitions). C'est la définition du fait que $a_0 \cdots a_n$ est reconnu par $\mathcal{A}$. $\{q_I\}^\perp$ est donc le langage reconnu par $\mathcal{A}$.

Soit $S \subseteq Q$. S est fini et s'écrit donc $\{q_0, \ldots, q_n\}$. Pour tout $q \in S$, on a $\{q\}^\perp$ rationnel puisque reconnu par l'automate $(\Sigma, Q, \delta, q, Q_F)$. Par conséquent,

$$S^\perp = \bigcap_{0 \le i \le n} \{q_i\}^\perp$$

est rationnel car il s'agit d'une intersection finie de langages rationnels. □

Une autre comparaison se trouve dans les transparents du cours de Paul-André Melliès (cf. [Mel10]) : afin d'introduire l'orthogonalité, pour une logique intuitionniste du second ordre, il utilise une analogie avec la géométrie algébrique et la topologie de Zariski.

3.3. Interprétation des formules et adéquation

On définit maintenant conjointement $|A| \subseteq \mathbb{T}$ et $||A|| \subseteq \mathbb{E}$.

Définition 15.
- $|A \to B| \stackrel{\text{def}}{=} ||A \to B||^{\perp}$ et $||A \to B|| \stackrel{\text{def}}{=} (|A| \cdot ||B||)^{\perp\perp} \stackrel{\text{def}}{=} \{t \cdot E \mid t \in |A| \text{ et } E \in ||B||\}^{\perp\perp}$,
- $|A \wedge B| \stackrel{\text{def}}{=} (|A| \times |B|)^{\perp\perp} \stackrel{\text{def}}{=} \{(t,u) \,\big|\, t \in |A| \text{ et } u \in |B|\}^{\perp\perp}$ et $||A \wedge B|| \stackrel{\text{def}}{=} |A \wedge B|^{\perp}$,
- $|A \vee B| \stackrel{\text{def}}{=} (|A| \uplus |B|)^{\perp\perp} \stackrel{\text{def}}{=} (\{(1,t) \mid t \in |A|\} \cup \{(2,t) \mid t \in |B|\})^{\perp\perp}$ et $||A \vee B|| \stackrel{\text{def}}{=} |A \vee B|^{\perp}$,
- $|\top| \stackrel{\text{def}}{=} \mathbb{T} = \emptyset^{\perp}$ et $||\top|| \stackrel{\text{def}}{=} |\top|^{\perp}$,
- $|\bot| \stackrel{\text{def}}{=} ||\bot||^{\perp}$ et $||\bot|| \stackrel{\text{def}}{=} \mathbb{E} = \emptyset^{\perp}$.

Indication pour lire les deux premières lignes (les autres se comprennent de façon similaire) :
- les programmes de type $A \to B$ doivent produire une interaction « correcte » vis-à-vis des piles $t \cdot E$ avec t de type A, et E acceptant une réponse de type B ;
- les contextes de type $A \wedge B$ sont ceux qui acceptent des paires (t,u) avec t de type A et u de type B. Un terme de type $A \wedge B$ est un programme qui produit une interaction « correcte » vis-à-vis de ces contextes-là.

Par définition, $|A|$ et $||A||$ sont polaires l'un de l'autre.

Théorème 16 (Lemme d'adéquation)**.** *Soit* $\Gamma = x_1 : A_1, \dots, x_n : A_n \stackrel{not.}{=} (\vec{x_i} : \vec{A_i})$.
- *Si* $\Gamma \vdash t : A$ *alors quels que soient* $t_1 \in |A_1|, \dots, t_n \in |A_n|$, *on a* $t[t_1/x_1, \dots, t_n/x_n] \in |A|$.
  - On note $[t_1/x_1, \dots, t_n/x_n]$ la substitution simultanée des t_i aux x_i.
  - On utilisera dans la suite la notation $t\left[\vec{t_i}/\vec{x_i}\right] \in |A|$ pour faire plus court.
- *Si* $\Gamma \mid E : A \vdash \star : B$ *alors quels que soient* $t_1 \in |A_1|, \dots, t_n \in |A_n|$ *et* $E' \in ||B||$, *on a* $E\left[\vec{t_i}/\vec{x_i}, E'/\star\right] \in ||A||$.
- *Si* $c : (\Gamma \vdash \star : B)$ *alors quels que soient* $t_1 \in |A_1|, \dots, t_n \in |A_n|$ *et* $E \in ||B||$, *on a* $c\left[\vec{t_i}/\vec{x_i}, E/\star\right] \in \perp\!\!\!\perp$.

Démonstration. Par induction structurelle sur la dérivation de t, E et c.
- Un certain nombre de cas sont évidents puisqu'ils découlent directement de la définition de $|A|$ ou de $||A||$. Il en va ainsi de l'introduction droite de $\wedge$ ($t = (u,v)$). En effet, si $u\left[\vec{t_i}/\vec{x_i}\right] \in |A|$ et $v\left[\vec{t_i}/\vec{x_i}\right] \in |B|$, alors $(u,v)\left[\vec{t_i}/\vec{x_i}\right] \in |A| \times |B| \subseteq |A \wedge B|$. Les cas des introductions de $\vee$ et de $\top$ à droite, ainsi que ceux des introductions de $\to$ et de $\bot$ à gauche, sont similairement évidents.
- Cas de $\overline{\Gamma', x : A \vdash x : A}$ avec $\Gamma', x : A = \Gamma$ et $x = t$. Par hypothèse, celui des t_i par lequel on substitue x appartient à $|A|$, donc $t\left[\vec{t_i}/\vec{x_i}\right] = t_i \in |A|$.
- Cas de $\overline{\Gamma \mid \star : A \vdash \star : A}$ avec $B = A$ et $E = \star$. Par hypothèse, $\star\left[\vec{t_i}/\vec{x_i}, E'/\star\right] = E' \in ||A||$.
- Cas de la coupure, avec $c = \langle t \mid E \rangle$: si $t\left[\vec{t_i}/\vec{x_i}\right] \in |A|$ et $E\left[\vec{t_i}/\vec{x_i}, E'/\star\right] \in ||A||$, alors :

$$t\left[\vec{t_i}/\vec{x_i}\right] \perp\!\!\!\perp E\left[\vec{t_i}/\vec{x_i}, E'/\star\right]$$

puisque $|A|$ et $||A||$ sont polaires. C'est-à-dire $c\left[\vec{t_i}/\vec{x_i}, E'/\star\right] \in \perp\!\!\!\perp$.

- Cas de l'adjoint : $t = c^*$. Soit $E \in ||A||$: il s'agit de montrer $c\left[\vec{t_i}/\vec{x_i}\right]^* \perp\!\!\!\perp E$. Or on a :

$$\left\langle c\left[\vec{t_i}/\vec{x_i}\right]^* \,\middle|\, E \right\rangle \succ c\left[\vec{t_i}/\vec{x_i}, E/\star\right] .$$

  Or $c\left[\vec{t_i}/\vec{x_i}, E/\star\right] \in \perp\!\!\!\perp$ par hypothèse d'induction. Comme $\perp\!\!\!\perp$ est $\mathbb{T},\mathbb{E}$-saturé et $E \in \mathbb{E}$ en particulier, on en déduit le résultat.
- Cas de l'introduction gauche de la conjonction $A \wedge B$. Soit $(t,u) \in |A| \times |B|$. Pour le contexte $\pi_1 \cdot E$, on a :

$$\left\langle (t,u) \,\middle|\, (\pi_1 \cdot E)\left[\vec{t_i}/\vec{x_i}, E'/\star\right] \right\rangle \succ \left\langle t \,\middle|\, E\left[\vec{t_i}/\vec{x_i}, E'/\star\right] \right\rangle .$$

  Or par hypothèse d'induction $E\left[\vec{t_i}/\vec{x_i}, E'/\star\right] \in ||A||$, donc $\left\langle t \,\middle|\, E\left[\vec{t_i}/\vec{x_i}, E'/\star\right] \right\rangle \in \perp\!\!\!\perp$ et donc puisque $t,u \in \mathbb{T}$, on a par $\mathbb{T},\mathbb{E}$-saturation $\left\langle (t,u) \,\middle|\, (\pi_1 \cdot E)\left[\vec{t_i}/\vec{x_i}, E'/\star\right] \right\rangle \in \perp\!\!\!\perp$. En conclusion :

$$(\pi_1 \cdot E)\left[\vec{t_i}/\vec{x_i}, E'/\star\right] \in (|A| \times |B|)^{\perp} = ||A \wedge B|| .$$

  Le cas de $\pi_2 \cdot E$ est similaire.
- Cas de l'introduction gauche de la disjonction $A \vee B$. Rappelons la règle qui est quelque peu inhabituelle car je voulais conserver une machine très élémentaire.

$$\frac{\Gamma \mid E : C \vdash \star : D \qquad \Gamma \vdash t : A \to C \qquad \Gamma \vdash u : B \to C}{\Gamma \mid (t|u) \cdot E : A \vee B \vdash \star : D}$$

  Soit alors $(i,v) \in |A| \uplus |B|$. On a donc $i = 1$ et $v \in |A|$, ou $i = 2$ et $v \in |B|$. Si $i = 1$, alors on a :

$$\left\langle (i,v) \,\middle|\, ((t|u) \cdot E)\left[\vec{t_i}/\vec{x_i}, E'/\star\right] \right\rangle \succ \left\langle t\left[\vec{t_i}/\vec{x_i}\right] \,\middle|\, v \cdot E\left[\vec{t_i}/\vec{x_i}, E'/\star\right] \right\rangle .$$

  Or par hypothèse d'induction $E\left[\vec{t_i}/\vec{x_i}, E'/\star\right] \in ||C||$ (le B de l'énoncé est ici D), donc :

$$v \cdot E\left[\vec{t_i}/\vec{x_i}, E'/\star\right] \in |A| \cdot ||C|| .$$

  Or $t\left[\vec{t_i}/\vec{x_i}\right] \in |A \to C| = (|A| \cdot ||C||)^{\perp}$ à nouveau par hypothèse d'induction, et donc $\left\langle t \,\middle|\, v \cdot E\left[\vec{t_i}/\vec{x_i}, E'/\star\right] \right\rangle \in \perp\!\!\!\perp$. Donc puisque $t\left[\vec{t_i}/\vec{x_i}\right], u\left[\vec{t_i}/\vec{x_i}\right] \in \mathbb{T}$, on a par $\mathbb{T},\mathbb{E}$-saturation $\left\langle v \,\middle|\, ((t|u) \cdot E)\left[\vec{t_i}/\vec{x_i}, E'/\star\right] \right\rangle$. D'où :

$$((t|u) \cdot E)\left[\vec{t_i}/\vec{x_i}, E'/\star\right] \in (|A| \uplus |B|)^{\perp} = ||A \vee B|| .$$

  Le cas $i = 2$ est similaire.
- Cas de l'introduction droite de $A \to B$:

$$\frac{\Gamma, x : A \vdash t : B}{\Gamma \vdash \lambda x.t : A \to B}$$

  Soit $u \cdot E \in |A| \cdot ||B||$. Il s'agit de montrer $(\lambda x.t)\left[\vec{t_i}/\vec{x_i}\right] \perp\!\!\!\perp u \cdot E$. Puisque les x_i sont distincts de x par hypothèse, et qu'on peut supposer que x n'apparaît pas dans les t_i (quitte à renommer x), on a donc :

$$\left\langle (\lambda x.t)\left[\vec{t_i}/\vec{x_i}\right] \,\middle|\, u \cdot E \right\rangle \succ \left\langle t\left[\vec{t_i}/\vec{x_i}, u/x\right] \,\middle|\, E \right\rangle .$$

  Or par hypothèse d'induction, on a $t\left[\vec{t_i}/\vec{x_i}, u/x\right] \in |B|$: on peut donc à nouveau déduire de la $\mathbb{T},\mathbb{E}$-saturation de $\perp\!\!\!\perp$ le résultat.

□

3.4. Applications

Proposition 17 (Cohérence)**.** *Il n'existe pas de terme clos t tel que $\vdash t : \bot$.*

Démonstration. On prend $\mathbb{T}$ et $\mathbb{E}$ les ensembles des termes et des contextes clos. Il suffit de prendre une structure de réalisabilité telle que $|\bot| = \emptyset$, par exemple $\perp\!\!\!\perp = \emptyset$, et de conclure avec le lemme d'adéquation. □

Remarquons tout de même que pour $\perp\!\!\!\perp = \emptyset$, on a :

$$|A| = \begin{cases} \mathbb{T} & \text{si } A \text{ est vrai,} \\ \emptyset & \text{si } A \text{ est faux,} \end{cases}$$

$$||A|| = \begin{cases} \emptyset & \text{si } A \text{ est vrai,} \\ \mathbb{E} & \text{si } A \text{ est faux.} \end{cases}$$

Par conséquent, la démonstration n'a pas grand intérêt ! On a seulement réussi à justifier les règles du système en « calculant » la vérité des formules dans un système plus puissant : on a déduit que si A est faux ou B est vrai, alors $A \to B$ est vrai, etc.

Mais Gentzen déjà s'intéressait moins au résultat de cohérence en lui-même qu'à la façon dont on l'obtient. En théorie de la démonstration, on s'intéresse au « comment » des preuves, et non seulement à la prouvabilité. C'est pourquoi les deux prochains résultats, qui donnent des propriétés des termes de preuve en utilisant une structure de réalisabilité non triviale, sont plus intéressants.

Théorème 18 (Propriété de la disjonction)**.** *Si $\vdash t : A \vee B$ alors il existe u tel que $\langle t \,|\, \star\rangle \succ \cdots \succ \langle (i,u) \,|\, \star\rangle$ pour $i = 1$ ou $i = 2$.*

Démonstration. On pose $\mathbb{T}$ et $\mathbb{E}$ les ensembles des termes et des contextes clos, ainsi que :

$$\perp\!\!\!\perp = \left\{ \langle t \,|\, E\rangle \in \mathbb{T} \times \mathbb{E} \,\middle|\, \exists u \in \mathbb{T}, \exists i \in \{1,2\}\, \langle t \,|\, E\rangle \succ \cdots \succ \langle (i,u) \,|\, \star\rangle \right\}.$$

On avait déjà remarqué que cela suffit à définir une structure de réalisabilité. Par adéquation, pour montrer $\langle t \,|\, \star\rangle \in \perp\!\!\!\perp$, il suffit de montrer $\star \in ||A \vee B||$. Or par définition de $\perp\!\!\!\perp$, $\star$ est orthogonal à tout $(i,u) \in |A| \uplus |B|$. D'où le résultat car $||A \vee B|| = (|A| \uplus |B|)^{\perp}$. □

Pour autant, que sait-on sur ce terme u ?

- En montrant une propriété syntaxique du calcul, la propriété de *réduction du sujet*, on obtient que u est une preuve de A si $i = 1$ et de B sinon. Cette propriété dit que si $c : (\Gamma \vdash \star : A \vee B)$ et $c \succ c'$, alors $c' : (\Gamma \vdash \star : A \vee B)$.
- Sans passer par une propriété syntaxique, en ayant plus de précisions sur A et B on peut raffiner le choix de $\perp\!\!\!\perp$ afin d'obtenir un résultat plus précis.

Définition 19. Une machine c est *normalisante* s'il existe c' avec $c \succ^* c'$ mais jamais $c' \succ c''$. Un terme t est normalisant si la machine $\langle t \,|\, \star\rangle$ l'est. Un contexte E est normalisant si pour toute variable x la machine $\langle x \,|\, E\rangle$ l'est.

Théorème 20 (Normalisation de tête)**.** *Si* $c : (x_1 : A_1, \ldots, x_n : A_n \vdash \star : A)$ *alors* c *se normalise.*

Démonstration. On pose $\mathbb{T}$ l'ensemble des termes normalisants, $\mathbb{E}$ l'ensemble des contextes normalisants, et :

$$\perp\!\!\!\perp = \{\langle t \mid E\rangle \in \mathbb{T} \times \mathbb{E} \,\big|\, \langle t \mid E\rangle \text{ est normalisant}\} .$$

S'il s'agit bien d'une structure de réalisabilité, alors par le lemme d'adéquation, pour montrer $c \in \perp\!\!\!\perp$ (le résultat), il suffit de montrer $x_1 \in |A_1|, \ldots, x_n \in |A_n|$ et $\star \in ||A||$. C'est une conséquence immédiate du fait que $||A_1||, \ldots, ||A_n||$ et $|A|$ sont constitués de contextes et de termes normalisants, c'est-à-dire tels que $\langle x_i \,|\, E\rangle$ et $\langle t \mid \star\rangle$ sont normalisants.

Montrons que $(\mathbb{T}, \mathbb{E}, \perp\!\!\!\perp)$ constitue bien une structure de réalisabilité. L'essentiel est de montrer que $\perp\!\!\!\perp$ est $\mathbb{T}, \mathbb{E}$-saturé. Or si $c \succ c'$ et que c' est normalisant, alors c est normalisant. Mais cela ne suffit pas : il est aussi nécessaire de montrer que c appartient à $\mathbb{T} \times \mathbb{E}$. Pour les cas $c = \langle (t_1, t_2) \,|\, \pi_i \cdot E\rangle$, $c = \langle (i,u) \,|\, (t_1|t_2) \cdot E\rangle$ et $c = \langle \lambda x.t \,|\, u \cdot E\rangle$ c'est évident. Pour $c = \langle c'^* \,|\, E\rangle$, on a par hypothèse $E \in \mathbb{E}$. Par ailleurs si c'^* ne normalise pas, alors c' non plus puisque $\langle c'^* \,|\, \star\rangle \succ c'$. Par conséquent $c'\,[E/\star]$ non plus, ce qui est contraire aux hypothèses. □

Remarque 21. La note se voulait introductive du point de vue des systèmes logiques considérés, et j'ai donc commencé par ne présenter qu'un calcul des prédicats intuitionnistes. On aurait pu se contenter, dans cette logique très simple, de démonstrations de nature combinatoire.

L'intérêt de la technique exposée ici est qu'on l'étend très facilement aux logiques du second ordre, au calcul des prédicats, à la logique classique, et fait partie d'un champ de recherche actif en logique mathématique et dans le milieu académique de la théorie des langages de programmation.

Cette technique est en fait une évolution d'une preuve de normalisation forte du système F due à Krivine, voir ci-dessous.

Remerciements Merci à l'audience du GdT pour sa participation, ainsi que pour avoir pointé un certain nombre de coquilles dans les précédentes versions. Merci à Pierre-Louis Curien pour des remarques sur cette note.

A. Addendum (Juin 2012) : Normalisation forte et comparaison avec la technique de la « paire adaptée »

A.1. Normalisation forte

La preuve s'adapte de la normalisation faible facilement pour prouver la normalisation *forte*. Cette propriété concerne la réduction $\succ$ étendue de façon contextuelle, que l'on note $\succcurlyeq$. On a donc $c \succcurlyeq c'$ si c' diffère de c par la réduction d'une sous-commande de c. On a juste besoin de quelques propriétés de « réécriture » qui concernent cette réduction contextuelle :

Proposition 22 (Weak standardisation)**.** *Si* $c\ (\succcurlyeq \setminus \succ)^*\ d \succ d'$ *alors il existe* c' *avec* $c \succ c' \succcurlyeq^* d'$.

On rappelle que r^* est la notation conventionnelle de la clôture transitive et réflexive de r.

Démonstration. Supposons que $c\ (\succcurlyeq \setminus \succ)^*\ d \succ d'$ en notant $c = \langle t \,|\, E\rangle$ et $d = \langle u \,|\, F\rangle$. Cela signifie que c se réduit en d par une succession de réductions sur les sous-commandes respectives de t et E. La possibilité de commuter provient du fait que les réductions principales que l'on a défini préservent les sous-commandes (dans le cas où $d \succ d'$ est une réduction principale) ; et que si $u = d''^*$, alors $t = c''^*$ avec $c'' \succcurlyeq d''$, donc $c \succ c''\,[E/\star] \succcurlyeq^* d''\,[E/\star] \succcurlyeq^* d''\,[F/\star] = d'$ (dans le cas où $d \succ d'$ est une réduction adjointe). □

Définition 23. Une commande c est fortement normalisante s'il n'existe pas de suite de réductions $\succcurlyeq$ infinie partant de c. On dit aussi que t et E sont fortement normalisants si les commandes $\langle t \,|\, \star\rangle$ et $\langle x \,|\, E\rangle$ (pour toute variable x) le sont.

Proposition 24. *Soit* $c = \langle t \,|\, E\rangle$ *une commande. Si aucune sous-commande de* t *ou de* E *n'admet de réduction* $\succcurlyeq$ *infinie, et si* $c \succ c'$ *avec* c' *fortement normalisant, alors* c *est fortement normalisant.*

Démonstration. Supposons que c admet une suite de réductions $\succcurlyeq$ infinie alors que par ailleurs $c \succ c'$ avec c' fortement normalisant. Supposons que la suite de réductions fait intervenir une réduction de tête ($\succ$) et notons $d \succ d'$ la première des réductions de tête. On a donc $c\ (\succcurlyeq \setminus \succ)^*\ d \succ d'$. D'après la proposition 22, il existe c'' avec $c \succ c'' \succcurlyeq^* d'$. Par déterminisme de $\succ$, on a alors $c' = c''$. C'est impossible, car c'' n'est pas fortement normalisant alors que c' l'est.

Par conséquent, la suite de réductions $\succcurlyeq$ infinie de c n'implique pas de réduction de tête. Chaque réduction concerne donc l'une des sous-commandes immédiates de t ou de E. Il y a une infinité de telles réductions, alors qu'il n'y a qu'un nombre fini de sous-commandes immédiates de t ou de E : l'une d'elle au moins n'est donc pas fortement normalisante. □

Proposition 25. *Aucune sous-commande de* t *(ou de* E*) n'admet de réduction* $\succcurlyeq$ *infinie si et seulement si* t *(resp.* E*) est fortement normalisant.*

Démonstration. ($\Leftarrow$) Immédiat. ($\Rightarrow$) Si t (resp. E) n'est pas fortement normalisant, alors $\langle t \,|\, \star\rangle$ (resp. $\langle x \,|\, E\rangle$) admet une suite de réductions $\succcurlyeq$ infinie. Si cette suite de réductions ne fait pas intervenir de réduction de tête, alors elle contient une suite de réductions infinie sur une sous-commande de t (resp. E). Sinon, par la proposition 22, on a la réduction adjointe $\langle t \,|\, \star\rangle \succ c$ avec $t = c^*$ et une suite de réductions infinie partant de c. D'où le résultat. □

En particulier, si t et E sont fortement normalisants, et $\langle t \,|\, E\rangle \succ c$ avec c fortement normalisant, alors $\langle t \,|\, E\rangle$ est fortement normalisant. Remarquer que le sens direct de la preuve est redevable de la définition des règles d'élimination par adjoint, qui permet par exemple de mettre en évidence que $\langle t \,|\, u \cdot \star\rangle$ est une sous-commande du terme $t\,u$!

Théorème 26 (Normalisation forte)**.** *Si* $c : (x_1 : A_1, \dots, x_n : A_n \vdash \star : A)$*, alors* c *est fortement normalisant.*

Démonstration. Similairement à la normalisation faible, il suffit de montrer que la structure ci-dessous définit une structure de réalisabilité :

- $\mathbb{T}$ l'ensemble des termes fortement normalisants ;
- $\mathbb{E}$ l'ensemble des contextes fortement normalisants ;
- $t \perp\!\!\!\perp E$ lorsque $t \in \mathbb{T}$, $E \in \mathbb{E}$ et $\langle t \,|\, E\rangle$ est fortement normalisant.

En effet, s'il s'agit bien d'une structure de réalisabilité, alors par le lemme d'adéquation, $c \in \perp\!\!\!\perp$ découle de $x_1 \in |A_1|, \dots, x_n \in |A_n|$ et de $\star \in ||A||$, ce qui est immédiat car on a par définition $\star \in \mathbb{T}^\perp$ et $x_i \in \mathbb{E}^\perp$ pour tout $i \leq n$.

$\mathbb{T}$ et $\mathbb{E}$ sont clos par les opérations positives : c'est immédiat avec la caractérisation de ces ensembles par la proposition 25. Montrons donc que $\perp\!\!\!\perp$ est $\mathbb{T}, \mathbb{E}$-saturé :

1. Si $E \in \mathbb{E}$ et $c\,[E/\star] \in \perp\!\!\!\perp$ alors $c^* \in \mathbb{T}$ par la proposition 25, puisque c est fortement normalisant (sinon $c\,[E/\star]$ ne le serait pas). A fortiori, $\langle c^* \,|\, E\rangle$ est fortement normalisant, les conditions de la proposition 24 étant réunis. Donc $\langle c^* \,|\, E\rangle \in \perp\!\!\!\perp$.
2. Si $t_1, t_2 \in \mathbb{T}$ et $\langle t_i \,|\, E\rangle \in \perp\!\!\!\perp$ pour $i \in \{1,2\}$. Alors $\langle (t_1, t_2) \,|\, \pi_i \cdot E\rangle \in \perp\!\!\!\perp$ provient de $(t_1, t_2) \in \mathbb{T}$ et de la proposition 24.
3. Si $t_1, t_2 \in \mathbb{T}$ et $\langle t_i \,|\, u \cdot E\rangle \in \perp\!\!\!\perp$ pour $i \in \{1,2\}$. Alors $u \in \mathbb{T}$ et $E \in \mathbb{E}$, et par conséquent $(i,u) \in \mathbb{T}$ et $(t_1|t_2)\cdot E \in \mathbb{E}$ aussi. Donc $\langle (i,u) \,|\, (t_1|t_2)\cdot E\rangle \in \perp\!\!\!\perp$ provient de la proposition 24.
4. Si $u \in \mathbb{T}$ et $\langle t\,[u/x] \,|\, E\rangle \in \perp\!\!\!\perp$. On a $t\,[u/x] \in \mathbb{T}$, donc $t \in \mathbb{T}$ aussi, et $\lambda x.t \in \mathbb{T}$ aussi. Par ailleurs $E \in \mathbb{E}$, donc $u \cdot E \in \mathbb{E}$. Donc $\langle \lambda x.t \,|\, u \cdot E\rangle \in \perp\!\!\!\perp$ provient de la proposition 24.

□

A.2. Comparaison avec la méthode de la « paire adaptée » de Krivine.

L'interprétation des types par orthogonalité, c'est « bien connu », est apparentée à la preuve de normalisation forte du Système F. Je montre que la technique est en effet similaire à la technique de la *paire adaptée* due à Krivine [Kri90, Chapitre 8, Section 3] (mais je ne connais pas le lien entre celle-ci et la preuve originale de Girard utilisant des candidats de réductibilité [Gir72]).

λ calcul oblige, le résultat concerne une notion de réduction sur les termes.

Définition 27. On note $t \rightsquigarrow t'$ lorsque $\langle t \,|\, \star\rangle \succcurlyeq^+ \langle t' \,|\, \star\rangle$ (où $\succcurlyeq^+$ est la clôture transitive de $\succcurlyeq$).

Remarquer que l'on a $(\lambda x.t)u\vec{v} \rightsquigarrow (t\,[u/x])\vec{v}$. (On rappelle la définition $t\,u = \langle t \,|\, u \cdot \star\rangle^*$ qui permet de vérifier à la main.)

Ici, je montre que la notion de structure de réalisabilité est, du point de vue de la réduction $\rightsquigarrow$, un cas particulier de la notion de *paire adaptée* due à Krivine [Kri90, Chapitre 8, Section

3]. Par ailleurs, lorsque les types sont interprétés par des comportements, les propriétés de cette paire adaptée utiles à la preuve proviennent de l'orthogonalité : on peut comprendre cela comme le fait que la présentation par orthogonalité permet de factoriser la preuve de normalisation forte à travers la preuve précédente que $\perp\!\!\!\perp$ est $\mathbb{T},\mathbb{E}$-saturé.

On peut donc voir la preuve de [Kri90] comme une troncature (dans le sens de la largeur, pas de la longueur) de la technique par orthogonalité du même Krivine, dans le sens où cette dernière généralise la première à une notion de réduction plus large.

On rappelle que la technique de Krivine pour démontrer la forte $\rightsquigarrow$-normalisation consiste à considérer la *paire adaptée* $(\mathcal{N}_0,\mathcal{N})$ avec $\mathcal{N}$ l'ensemble des termes fortement $\rightsquigarrow$-normalisables et $\mathcal{N}_0$ l'ensemble des termes de la forme $x\,t_1...t_n$ avec x une variable et $t_1,\dots,t_n\in\mathcal{N}$.

Les types sont interprétés dans des ensembles $\mathcal{X}$ avec $\mathcal{N}_0\subseteq\mathcal{X}\subseteq\mathcal{N}$: l'inclusion $\mathcal{N}_0\subseteq\mathcal{X}$ permet d'appliquer le lemme d'adéquation en substituant les variables par elles-mêmes, tandis que $\mathcal{X}\subseteq\mathcal{N}$ permet de conclure que les termes les termes typables sont dans $\mathcal{N}$, c'est-à-dire sont fortement $\rightsquigarrow$-normalisables. Il y a comme d'habitude une condition de saturation qui est demandée :

Définition 28. Un ensemble de termes $\mathcal{X}$ est $\mathcal{N}$-saturé si et seulement si :

$$u\in\mathcal{N}\text{ et }t\,[u/x]\,\vec{v}\in\mathcal{X}\Rightarrow(\lambda x.t)u\vec{v}\in\mathcal{X}\,.$$

Que la paire $(\mathcal{N}_0,\mathcal{N})$ soit adaptée correspond aux propriété suivantes :
- $\mathcal{N}$ est $\mathcal{N}$-saturé ;
- $\mathcal{N}_0\subseteq\mathcal{N}$; $\mathcal{N}_0\subseteq\mathcal{N}\to\mathcal{N}_0$; $\mathcal{N}_0\to\mathcal{N}\subseteq\mathcal{N}$;

où $\mathcal{X}\to\mathcal{Y}\stackrel{\text{def}}{=}\{t\mid\forall u\in\mathcal{X},\,tu\in\mathcal{Y}\}$.

La technique par orthogonalité revient à remplacer la paire adaptée $(\mathcal{N}_0,\mathcal{N})$ par la paire $(\mathbb{E}^\perp,\mathbb{T})$, et l'interprétation des types par des comportements ($\mathcal{X}=\mathcal{X}^{\perp\perp}$). En effet, le couple $(\mathbb{E}^\perp,\mathbb{T})$ a des propriétés très similaires :
- On a $\mathbb{T}\subseteq\mathcal{N}$: pour montrer la forte $\rightsquigarrow$-normalisation, il suffit de montrer la forte $\succcurlyeq$-normalisation.
- Si on pose $\mathbb{T}_0$ l'ensemble des termes de la forme $x\,t_1...t_n$ avec x une variable et $t_1,\dots,t_n\in\mathbb{T}$, alors $\mathbb{T}_0\subseteq\mathbb{E}^\perp$. En effet, si $x\,t_1...t_n\in\mathbb{T}_0$ et $E\in\mathbb{E}$, alors $\langle x\mid t_1\cdots t_n\cdot E\rangle\in\perp\!\!\!\perp$ par la proposition 25. Comme $\perp\!\!\!\perp$ est $\mathbb{T},\mathbb{E}$-saturé, on a $\langle x\,t_1...t_n\mid E\rangle\in\perp\!\!\!\perp$. Donc $x\,t_1...t_n\in\mathbb{E}^\perp$.
- La paire $(\mathbb{E}^\perp,\mathbb{T})$ est adaptée : en effet, $\mathbb{T}$ est $\mathbb{T}$-saturé (par la proposition 30 qui suit) ; et on vérifie que l'on a $\mathbb{E}^\perp\subseteq\mathbb{T}$; $\mathbb{E}^\perp\subseteq\mathbb{T}\to\mathbb{E}^\perp$; $\mathbb{E}^\perp\to\mathbb{T}\subseteq\mathbb{T}$.

Cependant, on n'a pas besoin de cette propriété de paire adaptée pour montrer, comme Krivine le fait pour la paire $(\mathcal{N}_0,\mathcal{N})$, que l'interprétation $\mathcal{X}$ d'un type vérifie $(\mathbb{T}_0\subseteq)\mathbb{E}^\perp\subseteq\mathcal{X}\subseteq\mathbb{T}$... puisque cela découle du fait que ce sont des comportements !

On a en particulier :

Proposition 29. *Pour tous ensembles de termes $\mathcal{X}\subseteq\mathbb{T}$ et de contextes $\mathcal{E}\subseteq\mathbb{E}$, on a :*

$$\mathcal{X}\to\mathcal{E}^\perp=(\mathcal{X}\cdot\mathcal{E})^\perp\,.$$

Démonstration. $t \in \mathcal{X} \to \mathcal{E}^\perp$ signifie $\forall u \in \mathcal{X}, \forall E \in \mathcal{E}, \langle tu \,|\, E\rangle \in \perp\!\!\!\perp$, tandis que $t \in (\mathcal{X} \cdot \mathcal{E})^\perp$ signifie $\forall u \in \mathcal{X}, \forall E \in \mathcal{E}, \langle t \,|\, u \cdot E\rangle \in \perp\!\!\!\perp$. Comme $\perp\!\!\!\perp$ est $\mathbb{T},\mathbb{E}$-saturé et que, s'agissant de la normalisation forte, $\perp\!\!\!\perp$ est stable par réduction, les deux conditions sont équivalentes. □

La dernière propriété qui intervient dans la technique de Krivine est que les types sont interprétés par des ensembles $\mathbb{T}$-saturés. C'est donc l'observation suivante qui permet l'articulation entre les deux techniques : la saturation de $\perp\!\!\!\perp$ pour $\succcurlyeq$ entraîne la saturation $\mathcal{X}$ pour $\rightsquigarrow$. (Qui est familier des travaux de Krivine reconnaît cette articulation comme une motivation pour interpréter les types par orthogonalité !)

Proposition 30. *Si $\mathcal{X} = \mathcal{X}^{\perp\perp}$ alors $\mathcal{X}$ est $\mathbb{T}$-saturé.*

On entend bien sûr par ensemble de termes $\mathbb{T}$-saturé un ensemble $\mathcal{X}$ tel que :

$$u \in \mathbb{T} \text{ et } t\,[u/x]\,\vec{v} \in \mathcal{X} \Rightarrow (\lambda x.t)u\vec{v} \in \mathcal{X}\,.$$

Démonstration. En effet, supposons $\mathcal{X} = \mathcal{X}^{\perp\perp}$. Soit $u \in \mathbb{T}$ et $t\,[u/x]\,\vec{v} \in \mathcal{X}$. On a $(\lambda x.t)u\vec{v} \rightsquigarrow t\,[u/x]\,\vec{v}$, donc $\langle (\lambda x.t)u\vec{v} \,\big|\, E\rangle \succcurlyeq^+ \langle t\,[u/x]\,\vec{v} \,\big|\, E\rangle$ pour tout $E \in \mathcal{X}^\perp$. On a $\langle t\,[u/x]\,\vec{v} \,\big|\, E\rangle \in \perp\!\!\!\perp$, donc aussi $\langle (\lambda x.t)u\vec{v} \,\big|\, E\rangle \in \perp\!\!\!\perp$ puisque $\perp\!\!\!\perp$ est $\mathbb{T},\mathbb{E}$-saturé. Donc $(\lambda x.t)u\vec{v} \in \mathcal{X}$. □

Pour être complet dans la comparaison, il nous faut introduire l'interprétation du second ordre bien que ce n'était pas présenté dans la note (je dois donc supposer le lecteur familier du système F). On interprète les variables du second ordre par des comportements, de sorte que $|X\,[\mathcal{X}/X]| = \mathcal{X}$ est un comportement. Cela signifie que l'on définit :

$$|\forall X\, A| \stackrel{\text{def}}{=} \bigcap \left\{ |A\,[\mathcal{X}/X]\,| \,\middle|\, \mathcal{X} = \mathcal{X}^{\perp\perp} \right\}$$

Comparer avec la définition dans le cadre de la paire adaptée :

$$|\forall X\, A| = \bigcap \left\{ |A\,[\mathcal{X}/X]\,| \,\middle|\, \mathcal{X} \text{ est } \mathcal{N}\text{-saturé}, \mathcal{N}_0 \subseteq \mathcal{X} \subseteq \mathcal{N} \right\}$$

L'une et l'autre des définitions nécessitent le *lemme de substitution* suivant :

$$|A\,[|B|/X]| = |A\,[B/X]|$$

Cela a un sens car un type est interprété, chez Krivine, par un ensemble $\mathcal{X}$ qui est $\mathcal{N}$-saturé et tel que $\mathcal{N}_0 \subseteq \mathcal{X} \subseteq \mathcal{N}$ (cela fait l'objet d'un lemme) ; et ici, un type est interprété par un ensemble de termes avec $\mathcal{X} = \mathcal{X}^{\perp\perp}$ (c'est immédiat par définition). Remarquer que cela implique qu'un type est encore interprété par un ensemble $\mathcal{X}$ qui est $\mathbb{T}$-saturé et tel que $\mathbb{E}^\perp \subseteq \mathcal{X} \subseteq \mathbb{T}$ (bien sûr, l'essentiel du travail a été fait en amont, lorsqu'on a prouvé que $\perp\!\!\!\perp$ est $\mathbb{T},\mathbb{E}$-saturé).

Le reste des deux preuves est identique : l'induction principale du lemme d'adéquation ; puis son application en se servant du fait que les variables libres sont toutes dans $\mathcal{N}_0/\mathbb{E}^\perp$.

Références

[CH00] Pierre-Louis Curien and Hugo Herbelin, *The duality of computation*, ACM SIGPLAN Notices **35** (2000), 233–243.

[Gir72] Jean-Yves Girard, *Interprétation fonctionnelle et élimination des coupures dans l'arithmétique d'ordre supérieur*, Ph.D. thesis, Université Paris VII, 1972.

[Gir06] ———, *Le point aveugle, cours de logique, tome i : Vers la perfection*, Vision des Sciences, Hermann, 2006.

[Kri90] Jean-Louis Krivine, *Lambda-calcul, types et modèles*, Masson, 1990, 175 p.

[Kri09] ———, *Realizability in classical logic*, Panoramas et Synthèses **27** (2009), 197–229.

[Mel10] Paul-André Melliès, *Logique du second ordre et réalisabilité*, Slides of the course "lambda-calcul et catégories", `http://www.pps.jussieu.fr/~mellies/mpri/mpri-ens/ens-mellies-cours-3.pdf`, 2010.

[Mun09] Guillaume Munch-Maccagnoni, *Focalisation and classical realisability*, Proc. CSL '09, LNCS, Springer-Verlag, 2009.