Rencontre LOCI au CIRM les 20 et 21 janvier 2011

Introduction à (une introduction à) la Ludique

Lionel Vaux

jeudi 20 janvier

La ludique provient d'une étude approfondie du caractère dynamique des preuves : l'interaction entre preuves et contre-preuves est au cœur du système, et la logique est reconstruite à partir de l'interaction.

L'artefact de base en ludique est la notion de dessin :
- on efface la distinction syntaxe/sémantique, en élargissant l'espace des preuves (on peut construire des preuves non correctes) ;
- on se concentre sur la structure calculatoire des preuves, en oubliant le « nom » des formules.

Un dessin est en quelque sorte un squelette de preuve : un tel squelette représente les manipulations locales que les règles de déduction font subir aux formules.

1 Prélude : Cohérence et complétude de LK

En étudiant le cas de la logique classique, on peut à la fois cerner le caractère logique de l'interaction et la manière dont on peut donner un statut sémantique à la syntaxe en élargissant l'espace des preuves.

En calcul des séquents, les jugements sont de la forme $\Gamma \vdash \Delta$ ou Γ et Δ sont des familles finies de formules : $A_1, \ldots, A_n \vdash B_1, \ldots, B_p$ dénote intuitivement $A_1 \wedge \cdots \wedge A_n \Rightarrow B_1 \vee \cdots \vee B_p$. On peut considérer un système monolatère à droite, un tel séquent étant classiquement équivalent à $\vdash \neg A_1, \ldots, \neg A_n, B_1, \ldots, B_p$. De plus, modulo les lois de Morgan, on peut restreindre les formules aux connecteurs $\wedge$ et $\vee$, en reportant les négations sur les atomes. On peut alors donner les règles de preuve :

$$\frac{}{\vdash \Gamma, \neg X, X}\,(\mathrm{ax}) \qquad \frac{\vdash \Gamma, \Delta, A \qquad \vdash \Gamma, \Delta', \neg A}{\vdash \Gamma, \Delta, \Delta'}\,(\mathrm{cut})$$

$$\frac{\vdash \Gamma, A \qquad \vdash \Gamma, B}{\vdash \Gamma, A \wedge B}\,(\wedge) \qquad \frac{\vdash \Gamma, A, B}{\vdash \Gamma, A \vee B}\,(\vee)\ .$$

Le théorème d'élimination des coupures établit que la règle (cut) est logiquement superflue : toute preuve avec (cut) peut être transformée en une preuve sans (cut) de même conclusion. [1] On en déduit en particulier la cohérence du calcul : le séquent vide $\vdash$ (absurde) n'est pas prouvable sans coupure.

Pour établir que le système sans coupure est complet pour le calcul booléen, une méthode consiste à remplacer les axiomes (ax) par une règle de terminaison arbitraire :

$$\frac{}{\vdash \Gamma_0, \neg\Delta_0}\ (\text{stop})$$

où les formules de Γ_0 et Δ_0 sont atomiques.

Remarquons qu'on a choisi les règles ($\wedge$) et ($\vee$) de sorte que la conclusion d'une règle soit équivalente à (la conjonction de) ses prémisses. Alors, partant de n'importe quel séquent $\vdash \Gamma$, on peut rechercher une preuve par application des règles ($\wedge$) et ($\vee$) jusqu'à épuisement des connecteurs, et en terminant par des instances de (stop). On obtient que $\vdash \Gamma$ est équivalent à la conjonction des instances de (stop) ainsi générées. On a alors deux possibilités :

- si toutes ces instances sont des axiomes (Γ_0 et Δ_0 ont une formule commune) alors Γ est vrai et prouvable (puisqu'on a une « vraie » preuve) ;
- sinon, l'une des instances de (stop) est falsifiable, ce qui donne un choix de vérité qui falsifie également Γ.

2 Calcul hyperséquentialisé

2.1 Logique linéaire

La logique linéaire décompose la logique intuitionniste à travers la fameuse traduction $A \Rightarrow B = {!A} \multimap B$. L'implication linéaire ($\multimap$) remplace la notion usuelle d'implication logique par celle d'une transformation : l'hypothèse est « consommée » pour produire la conclusion. La modalité exponentielle (!) compense cette sensibilité aux ressources en dénotant la réplicabilité de l'hypothèse. Les règles structurelles (contraction = duplication ; affaiblissement = effacement) ne sont en général disponibles que sur les formules exponentielles. En particulier le nombre d'occurrences d'une formule dans un séquent devient une donnée pertinente. Dans ce cadre, deux notions de conjonction ($\otimes$ et &) et deux notions de disjonction ($⅋$ et $\oplus$) coexistent sans être confondues : $\otimes$ et $⅋$ sont des connecteurs **multiplicatifs**, alors que & et $\oplus$ sont **additifs**. Cette terminologie est relative à la gestion du contexte dans les règles :

$$\frac{\vdash \Gamma, A \qquad \vdash \Delta, B}{\vdash \Gamma, \Delta, A \otimes B}\ (\otimes) \qquad \frac{\vdash \Gamma, A \qquad \vdash \Gamma, B}{\vdash \Gamma, A \,\&\, B}\ (\&)$$

$$\frac{\vdash \Gamma, A, B}{\vdash \Gamma, A ⅋ B}\ (⅋) \qquad \frac{\vdash \Gamma, A_i}{\vdash \Gamma, A_1 \oplus A_2}\ (\oplus_i)\ .$$

L'axiome et la coupure sont de nature multiplicative :

1. On a choisi la forme particulière des règles (ax) et (cut) pour conserver ce résultat, sans avoir à introduire de règles structurelles (contraction et affaiblissement) explicites. Ce détail technique n'est pertinent que parce qu'on a une autre idée en tête : la preuve de complétude qui suite.

$$\frac{}{\vdash X^\perp, X}\,(\mathrm{ax})\ , \qquad \frac{\vdash \Gamma, A \qquad \vdash A^\perp, \Delta}{\vdash \Gamma, \Delta}\,(\mathrm{cut})\ .$$

Nous travaillerons dans ce fragment multiplicatif-additif de la logique linéaire (donc sans exponentielle) qui est à la base de la ludique.

2.2 Réversibilité et focalisation

Un connecteur $\bullet$ est dit **réversible** si, pour tout séquent prouvable $\vdash A \bullet B, \Gamma$, il existe une preuve se terminant par l'introduction de $A \bullet B$.

Les connecteurs $\parr$ et & sont réversibles : des conclusions respectives des règles ($\parr$) et (&), on peut déduire chacune des prémisses, et donc appliquer de nouveau la règle correspondante.[2] En particulier, dans le cadre d'une recherche de preuve, on peut toujours commencer par appliquer les règles des connecteurs réversibles, sans perdre de pouvoir de preuve.

Les connecteurs $\otimes$ et $\oplus$ *ne sont pas* réversibles (contre-exemples : $\vdash A \otimes B, A^\perp \parr B^\perp$ et $\vdash A \oplus B, A^\perp \,\&\, B^\perp$). On peut pourtant établir une sorte d'associativité pour les règles de ces connecteurs, et ainsi limiter très fortement l'espace à explorer lors de la recherche de preuves.

Plus précisément, on organise les connecteurs en deux classes : les connecteurs **négatifs** sont les connecteurs réversibles $\parr$ et & ; les connecteurs **positifs** sont leurs duaux $\otimes$ et &. Et on organise les formules en fonction de la polarité de leur connecteur principal. On définit alors la propriété de focalisation comme suit. Une preuve (sans coupure) est dite **focalisée** si :

– chaque fois qu'on attaque une formule, on continue à décomposer ses sous-formules tant que celles-ci sont de la même polarité ;
– si un séquent contient des formules négatives, on commence par décomposer celles-ci.

On obtient alors que tout séquent prouvable admet une preuve focalisée. Pour ces preuves, lors d'une phase négative, l'ensemble des conclusions négatives forme un bloc, attaqué d'un seul tenant. Par contraste, dans une phase positive, on commence par choisir la formule positive sur laquelle on *focalise* la tentative de preuve.

2.3 Connecteurs synthétiques

La focalisation permet de structurer les preuves de la logique linéaire en une alternance de phases positives et négatives. Mieux, on peut voir ces phases comme la décomposition à grain fin de règles de déduction associées à des connecteurs **synthétiques** : un connecteur positif (généralisation de $\otimes$ et $\oplus$) et un connecteur négatif (généralisation de $\parr$ et &).

Ceci est rendu possible par les isomorphismes de distributivité $(A \otimes B) \oplus C \simeq A \otimes C \oplus B \otimes C$ et $(A \parr B) \,\&\, C \simeq A \parr C \,\&\, B \parr C$ (le mot isomorphisme n'est pas galvaudé : on a bien des couples de preuves réciproques modulo élimination des coupures).

2. Remarquons en passant que ces règles sont celles qu'on a données pour $\wedge$ et $\vee$, dans la section précédente : on les avait justement choisies pour cette propriété.

Par convention, on note P les formules positives ou les atomes, et N les formules négatives ou les négations d'atomes. On considèrera alors les formules générées à partir des atomes et de leurs négations, par l'application des connecteurs synthétiques : $\bigoplus_{i\in I}\bigotimes_{j\in J_i} N_{i,j}$ et $\&_{i\in I}\bigparr_{j\in J_i} P_{i,j}$.

On obtient les règles synthétiques :

$$\frac{(\vdash N_{i_0,j},\Gamma_j)_{j\in J_{i_0}} \qquad \Gamma=\sum_{j\in J_{i_0}}\Gamma_j}{\vdash \bigoplus_{i\in I}\bigotimes_{j\in J_i} N_{i,j},\Gamma}\,(+,i_0) \qquad \frac{\left(\vdash (P_{i,j})_{j\in J_i},\Gamma\right)_{i\in I}}{\vdash \&_{i\in I}\bigparr_{j\in J_i} P_{i,j},\Gamma}\,(-)$$

dans lesquelles le contexte Γ ne contient que des formules positives ou des atomes (éventuellement niés).

Une preuve dans ce nouveau système est une alternance entre règles positives et négatives, avec une structure proche des stratégies en théorie des jeux : les règles négatives représentent des questions, les règles positives étant les réponses (choix). Les coupures sont le lieu du dialogue entre deux stratégies.

2.4 Élimination des coupures

Nous n'avons pas donné les règles d'élimination des coupures en logique linéaire : nous le faisons pour le système hyperséquentialisé.

$$\frac{\dfrac{\begin{pmatrix}\pi_j\\ \vdots\\ \vdash P^{\perp}_{i_0,j},\Gamma_j\end{pmatrix}_{j\in J_{i_0}}}{\vdash \bigoplus_{i\in I}\bigotimes_{j\in J_i} P^{\perp}_{i,j},\Gamma}\,(+,i_0) \qquad \dfrac{\begin{pmatrix}\rho_{i,j}\\ \vdots\\ \vdash (P_{i,j})_{j\in J_i},\Delta\end{pmatrix}_{i\in I}}{\vdash \&_{i\in I}\bigparr_{j\in J_i} P_{i,j},\Delta}\,(-)}{\vdash \Gamma,\Delta}\,(\mathrm{cut})$$

$$\rightsquigarrow \quad \frac{\begin{pmatrix}\pi_j\\ \vdots\\ \vdash P^{\perp}_{i_0,j},\Gamma_j\end{pmatrix}_{j\in J_{i_0}} \qquad \begin{matrix}\rho_{i_0,j}\\ \vdots\\ \vdash (P_{i_0,j})_{j\in J_{i_0}},\Delta\end{matrix}}{\vdash (\Gamma_j)_{j\in J_{i_0}},\Delta}\,(\mathrm{cut})\times\#J_{i_0}$$

Intuitivement, la règle positive choisit l'une des alternatives proposées par la règle négative. Les sous-formules sont alors mises en correspondance, la polarité étant échangée dans ces nouvelles coupures.

3 Ludique

On a déjà souligné l'intérêt qu'il pouvait y avoir à enrichir le système de preuves « incorrectes ». En ludique, cette idée est fondamentale car ce système place le dialogue entre preuves et contre-preuves au cœur de la logique : la coupure confronte le défenseur de $\vdash A$ au tenant de $\vdash A^{\perp}$. Pour que ce dialogue ait lieu dans le cas où A est effectivement prouvable, il faut bien introduire ce qui tiendra lieu de « preuve » de $A^{\perp}$. L'essentiel est que les deux parties s'accordent pour déclarer un gagnant.

3.1 Dessins

La parapreuve la plus élémentaire est très similaire au (stop) de notre introduction : le **daimon** $\maltese$ est une règle positive qui symbolise l'abandon d'une stratégie. Plus précisément, on aurait

$$\frac{}{\vdash \Gamma}\ (\maltese)$$

où Γ ne contient que des formules positives. Cette « preuve » affirme défendre Γ, mais ne supporte pas qu'on lui demande d'argumenter : en cas de question (*i.e.* de coupure), le daimon abandonne et reconnaît son échec.

Par ailleurs, cette démarche impose l'élaboration d'un système sans *a priori* logique : on « oublie » le sens des formules pour n'en conserver que la structure, c'est-à-dire l'organisation en sous-formules. Plus précisément, les règles de la ludique agissent sur des adresses, *i.e.* des suites d'indices qui représentent les positions des sous-formules : on appelle **adresse** (ou lieu, ou *locus*) toute suite finie d'entiers naturels. Une sous-adresse de ξ est une adresse de la forme $\xi\xi'$. Autrement dit, une adresse est préfixe de ses sous-adresses. Les séquents de la ludique sont de la forme $\Lambda \vdash \Delta$ où Λ et Δ sont des ensembles finis d'adresses deux-à-deux disjointes, et Λ a au plus un élément : un tel séquent est dit **positif** si $\Lambda = \emptyset$ et **négatif** sinon (*cf.* la forme des séquents dans les preuves focalisées).

On aboutit enfin à la notion de dessin (*design*), qui abstrait celle de preuve. Un **dessin** est un arbre formé à partir des règles suivantes :

- daimon :

$$\frac{}{\vdash \Delta}$$

- règle positive :

$$\frac{(\xi i \vdash \Delta_i)_{i\in I}}{\vdash \xi, \Delta}\ (+,\xi,I)$$

 où I est un ensemble fini d'entiers et les Δ_i sont des parties deux-à-deux disjointes de Δ ;
- règle négative :

$$\frac{\big(\vdash (\xi i)_{i\in I}\,, \Delta_I\big)_{I\in\mathcal{N}}}{\xi \vdash \Delta}\ (-,\xi,\mathcal{N})$$

 où $\mathcal{N}$ est un ensemble de parties finies de $\mathbf{N}$ et chaque Δ_I est une partie de Δ.

Remarquons que chaque règle de connecteur synthétique correspond à une instance de ces règles abstraites, mais que la réciproque n'est pas vraie : le schéma ci-dessus est plus général. En particulier, la ludique admet les affaiblissements et les connecteurs d'arité quelconque ($\mathcal{N}$ peut être infini).

Par ailleurs, les dessins ne sont pas supposés être des arbres finis. En particulier, un dessin qui ne fait pas usage du démon est nécessairement infini. L'exemple le plus élementaire de tel dessein est le *fax* :

$$\dfrac{\cdots \qquad \dfrac{\cdots \qquad \overset{\vdots}{\xi' i \vdash \xi i} \qquad \cdots}{\vdash \xi', (\xi i)_{i\in I}}\ (+,\xi',I) \qquad \cdots}{\xi \vdash \xi'}\ (-,\xi,\mathfrak{P}_f(\mathbf{N}))$$

Ce dessin est en quelque sorte la représentation de l'axiome en ludique : sans formules atomiques, on construit une η-expansion généralisée.

Note : dessins et desseins. Les dessins ont une structure assez fortement séquentielle. Notamment, deux dessins qui ne diffèrent que par la répartition des contextes lors de l'application des règles sont distincts alors qu'ils représentent essentiellement les mêmes « stratégies ». Il est possible de lever cette ambiguïté, en introduisant des *desseins*, lesquels sont intuitivement obtenus en ne conservant que les étiquettes des règles : on obtient une structure très proche des celles des stratégies en sémantique des jeux, sur lesquelles on peut définir également définir une notion de normalisation, et dont les propriétés sont importantes pour développer toute la théorie ludique. Nous nous contenterons cependant des dessins, dans un souci de simplicité (et donc sans rentrer dans les détails de la théorie).

3.2 Normalisation

On appelle **réseau de coupures** tout ensemble fini non vide de dessins tel que :
- les adresses apparaissant dans les conclusions des dessins sont deux-à-deux disjointes ou identiques (pas de préfixe) ;
- chaque adresse apparaît dans au plus deux conclusions et, dans ce cas, une fois à gauche et une fois à droite du signe $\vdash$: une telle adresse est une **coupure** ;
- le graphe dont les sommets sont les conclusions et les arêtes les coupures est connexe et acyclique.

En particulier, il y a exactement un dessin qui ne contient pas de coupure à gauche : sa conclusion est le **séquent principal** du réseau, et sa dernière règle, la **règle principale**.

On se contente de traiter le cas où toutes les adresses de conclusions sont des coupures (réseau clos). Le dessin principal D est alors nécessairement positif. On a trois cas :
- La règle principale est (✠) : la normalisation est terminée et le résultat est réduit au daimon.
- La règle principale est $(+, \xi, I)$: alors ξ est une coupure, apparaît à gauche dans la conclusion d'un autre dessin E (le **dessin adjoint**), dont la dernière règle est nécessairement $(-, \xi, \mathcal{N})$. On distingue deux sous-cas :
  - si $I \notin \mathcal{N}$, la normalisation échoue ;
  - sinon, pour tout $i \in I$, on considère le sous-dessin D' (négatif) du dessein principal, de conclusion $(\xi i \vdash \cdots)$ et le sous-dessein E' (positif) du dessein adjoint, de conclusion $(\vdash \xi I, \cdots)$: on remplace alors D par D' et E par E'. La composante connexe de E' forme alors un nouveau réseau, qu'on normalise.

Le cas général (non clos) fait intervenir une notion de commutation de règles que nous ne développerons pas.

Pour comprendre pleinement ce qui est en jeu dans cette procédure, il faudrait avoir en tête l'élimination des coupures dans les réseaux de preuve. On peut en tout cas relier le dernier cas à l'élimination des coupures dans le calcul hyperséquentialisé, et se convaincre que les deux autres cas sont des formes distinctes de terminaison : le daimon est une défaite reconnue, tandis que l'échec de la normalisation dénote l'impossibilité pour deux desseins de « se parler ». Enfin, un dernier cas se présente : celui où la forme normale n'est pas définie, la définition récursive de la normalisation n'étant pas nécessairement bien fondée. Là encore, c'est l'échec de l'interaction, cette fois parce que le dialogue se poursuit sans fin. Ainsi, deux issues sont possibles pour la normalisation : le daimon (reconnaissance d'un résultat) et l'échec.

3.3 Comportements

Pour conclure cette courte introduction, nous allons donner une idée de la reconstruction des connecteurs logiques à partir de l'interaction. Celle-ci est basée sur la notion élémentaire d'orthogonalité.

Soit D un dessin de conclusion $\Lambda \vdash \Delta$ et soient des dessins E_ξ de conclusion $\xi \vdash$ si $\xi \in \Delta$ et $\vdash \xi$ si $\xi \in \Lambda$. On dira que D est **orthogonal** à la famille (E_ξ) si la normalisation du réseau $\{D\} \cup \{E_\xi\}_\xi$ mène au daimon.

Si $\mathbf{D}$ est un ensemble de dessins (partageant la même conclusion), on note $\mathbf{D}^{\perp\perp}$ son **bidual** : l'ensemble des dessins (de même conclusion), qui sont orthogonaux aux familles orthogonales à tous les dessins de $\mathbf{D}$. Lorsque un tel ensemble est égal à son bidual, on dit que c'est un **comportement** : les comportements sont l'équivalent ludique des formules.

Une intersection quelconque de comportements (de même conclusion) reste un comportement : $\bigcap_i \mathbf{D}_i$. Par contre une union n'est pas forcément close par bidual : on note $\bigsqcup \mathbf{D}_i = (\bigcup \mathbf{D}_i)^{\perp\perp}$. Lorsque les comportements positifs $\mathbf{D}_1$ et $\mathbf{D}_2$ sont disjoints (intuitivement : les indices de leurs premières règles diffèrent) on note $\mathbf{D}_1 \oplus \mathbf{D}_2 = \mathbf{D}_2 \sqcup \mathbf{D}_2$. Dans ce cas, on a en fait : $\mathbf{D}_1 \sqcup \mathbf{D}_2 = \mathbf{D}_2 \cup \mathbf{D}_2$. De la même manière, on note $\mathbf{D}_1 \,\&\, \mathbf{D}_2 = \mathbf{D}_2 \cap \mathbf{D}_2$ pour des comportements négatifs véifiant une condition duale. On retrouve ainsi les connecteurs additifs de la logique linéaire, mais dans une version *localisée* : on pose une condition sur les adresses pour pouvoir les définir.

Retrouver une version totale des additifs amène à développer une notion de délocalisation sur laquelle nous ne nous étendrons pas. Nous ne présenterons pas non plus les multiplicatifs ici : il existe en fait plusieurs manières de les construire, à partir d'une manipulation élémentaire sur les dessins (intuitivement, pour le tenseur ($\otimes$), on met en parallèle les premières actions). Pour conclure, indiquons que les preuves « correctes » seront les dessins vérifiant certaines propriétés, en particulier celle de ne pas recourir au daimon.

Références

[1] Jean-Yves Girard. Locus Solum : From the rules of logic to the logic of rules. *Mathematical Structures in Computer Science*, 11(3) :301–506, 2001.

[2] Thomas Seiller. La Ludique. Mémoire du Master 1 Logique, Philosophie, Histoire et Sociologie des Sciences, Université Paris 1 Panthéon Sorbonne, 2007.