Chapitre 1

Induction par rêgles

La notion de rêgle est très simple, mais elle offre un outil très puissant qui va nous permettre de modéliser plein de concepts utiles en informatique.

1.1 Notion de rêgles

Définition 1. On appelle rêgle sur A un élément $(h, B) \in A \times 2^A$, et on l'écrit habituellement $h \leftarrow B$

h est la tête (*head* en anglais) de la rêgle, B le corps (*body*). Lorsque B est un ensemble fini, on dit que la rêgle est *finitaire*. On note parfois une rêgle sous la forme d'une fraction du style "rêgle d'inférence" :

$$\frac{\{b_1, b_2, \ldots, b_n\}}{h}$$

On utilise souvent une notation allégée :

$h \leftarrow$	au lieu de	$h \leftarrow \emptyset$
$h \leftarrow b_1, \ldots, b_n$		$h \leftarrow \{b_1, \ldots, b_n\}$

mais attention à l'ambiguité pour $h \leftarrow b_1$ quand b_1 est aussi un ensemble.

1.2 Examples d'applications possibles

Nous donnons maintenant des examples pour illustrer la variété des applications possibles des rêgles, mais brièvement et sans entrer dans les détails.

Logique. on peut représenter les règles d'inférence comme *Modus Ponens* :

$$\frac{P \Rightarrow Q,\ P}{Q} \qquad\qquad Q \leftarrow \{P \Rightarrow Q,\ P\}$$

Graphe orienté.

$$\begin{tikzcd}
a \arrow[r] \arrow[d] & b \arrow[r] \arrow[l, bend left] & d \\
c & & \\
e \arrow[u] \arrow[rr] & & f \arrow[uu]
\end{tikzcd}
\qquad\qquad
\begin{array}{ll}
b \leftarrow a & c \leftarrow e \\
c \leftarrow a & f \leftarrow e \\
a \leftarrow b & d \leftarrow f \\
d \leftarrow b &
\end{array}$$

Questions fréquentes : quels sont les successeurs immédiats d'un nœud. Quels nœuds peut-on atteindre à partir d'un nœud ou un ensemble de nœuds donnés (fermeture transitive) ?

Relation binaire.

$$x\ R\ y \qquad\qquad y \leftarrow x$$

On parle d'ailleurs du graphe d'une relation. Fermeture transitive $x\ R^+\ y$ et réflexive transitive $x\ R^*\ y$.

Structures de données. Les listes d'entiers :

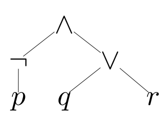
```
type intList = nil | int::intList
```

$$\begin{array}{r}
\texttt{nil} \leftarrow \\
x :: \ell \leftarrow \ell \quad \forall x \in \mathbb{N}
\end{array}$$

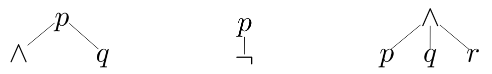
Syntaxe abstraite. Les formules de la logique propositionelle. Etant donné un ensemble V de variables propositionelles :

$$F \quad ::= \quad p \mid F \wedge F \mid F \vee F \mid \neg F \quad (p \in V) \qquad\qquad
\begin{array}{r}
p \leftarrow \emptyset \quad (\forall p \in V) \\
F_1 \wedge F_2 \leftarrow \{F_1, F_2\} \\
F_1 \vee F_2 \leftarrow \{F_1, F_2\} \\
\neg F \leftarrow \{F\}
\end{array}$$

Pour mieux comprendre ce qu'on vient d'écrire ci-dessus, on adopte le point de vue qu'une formule de la logique propositionelle est un arbre décoré par des symboles :

$\neg p \ \wedge\ (q \vee r) \qquad \equiv$ 

Parmis tous les arbres décorés qu'il est possible d'écrire, il y en a qui ne correspondent pas à des formules bien formées :



Les schémas tels que $F_1 \wedge F_2 \;\leftarrow\; \{F_1, F_2\}$ seront instantiés pour tous les arbres décorés F_1 et F_2, qu'ils correspondent à des formules bien formées ou non. On aura ainsi une infinité de règles : les règles ne correspondant pas à des formules bien formées ne seront simplement jamais applicables.

Grammaires hors-contexte. Considérons un alphabet A et la grammaire hors-contexte des palindromes sur A :

$$
\begin{array}{lll}
S \rightarrow \epsilon & & \\
S \rightarrow a & \forall a \in A & \\
S \rightarrow a\ S\ a & \forall a \in A &
\end{array}
\qquad\qquad
\begin{array}{rl}
\epsilon \leftarrow \emptyset & \\
a \leftarrow \emptyset & \forall a \in A \\
a\ m\ a \leftarrow m & \forall a \in A,\ m \in A^*
\end{array}
$$

Fonctions récursives. Considérons la fonction factorielle :

$$
\begin{array}{ll}
\mathrm{fact}(0) & = 1 \\
\mathrm{fact}(n+1) & = (n+1) * \mathrm{fact}(n)
\end{array}
$$

qu'on modélise par le système de rêgles suivant :

$$
\begin{array}{rl}
(0,1) \leftarrow \emptyset & \\
(n+1, (n+1) * x) \leftarrow (n, x) & \qquad \forall n, x \in \mathbb{N}
\end{array}
$$

Les entiers. On considère la construction des entiers proposée par von Neumann :

$$
\begin{array}{r}
0 \equiv \emptyset \\
x^+ \equiv x \cup \{x\}
\end{array}
\qquad\qquad
\begin{array}{r}
\emptyset \leftarrow \emptyset \\
x \cup \{x\} \leftarrow \{x\}
\end{array}
$$

Ce qui nous donne :

$$
\begin{array}{l}
0 = \emptyset \\
1 = \{\emptyset\} \\
2 = \{\emptyset, \{\emptyset\}\} \\
3 = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} \\
\quad\vdots
\end{array}
$$

Mais les entiers ainsi définis ont des propriétés additionelles inattendues : $n \in n+1$ et $n \subseteq n+1$

On écrit B^A pour l'ensemble des fonctions de A dans B. Un sous ensemble de A peut être décrit par sa fonction caractéristique $A \to \{0, 1\}$. Donc on peut identifier l'ensemble des parties de A avec l'ensemble des fonctions caractéristiques sur A, c'est à dire $\{0, 1\}^A$. Or, dans la représentation de von Neumann, $\{0, 1\}$ c'est 2. D'ou la notation 2^A pour l'ensemble des parties de A.

1.3 Image d'un ensemble par des rêgles

Soit $\mathcal{R}$ un ensemble de rêgles, fini ou infini.

$$\mathcal{H}(\mathcal{R}) = \{h \mid \exists h \leftarrow B \in \mathcal{R}\} \qquad \text{têtes de rêgles}$$
$$\mathcal{B}(\mathcal{R}) = \{B \mid \exists h \leftarrow B \in \mathcal{R}\} \qquad \text{corps de rêgles}$$

On notera alors $\cup\mathcal{B}(\mathcal{R})$ la réunion de tous les corps de rêgles.

Définition 2 (image)**.** Soit $\mathcal{R}$ un ensemble de rêgles et X un ensemble quelconque. On note $\mathsf{T}_{\mathcal{R}}(X)$ l'image de X par R définie par :

$$\mathsf{T}_{\mathcal{R}}(X) = \{h \mid \exists (h \leftarrow B) \in \mathcal{R},\ B \subseteq X\}$$

Dans un graphe orienté, $\mathsf{T}_{\mathcal{R}}(X)$ est l'ensemble des successeurs immédiats des nœuds X. En logique, $\mathsf{T}_{\mathcal{R}}(X)$ est l'ensemble des déductions immédiates sur la base des prémisses X.

- $\mathsf{T}_{\mathcal{R}}(\emptyset)$ est l'ensemble des têtes de faits
- $\mathsf{T}_{\mathcal{R}}(\emptyset) = \emptyset$ ssi il n'y a aucun fait
- $\mathsf{T}_{\mathcal{R}}(\cup\mathcal{B}(\mathcal{R})) = \mathcal{H}(\mathcal{R})$

Théorème 1. $\mathsf{T}_{\mathcal{R}}$ *est une fonction croissante :*

$$X \subseteq Y \quad \Rightarrow \quad \mathsf{T}_{\mathcal{R}}(X) \subseteq \mathsf{T}_{\mathcal{R}}(Y)$$

Preuve. $\forall (h \leftarrow B) \in \mathcal{R}$, si $B \subseteq X$, alors $B \subseteq X \subseteq Y$. Donc si $h \in \mathsf{T}_{\mathcal{R}}(X)$ alors $h \in \mathsf{T}_{\mathcal{R}}(Y)$ □

Définition 3 (fermé)**.** Un ensemble X est dit fermé, ou encore clos ou stable, pour $\mathcal{R}$ ssi $\mathsf{T}_{\mathcal{R}}(X) \subseteq X$

Autrement dit, par $\mathsf{T}_{\mathcal{R}}$ on ne sort pas de X ; ou encore : on ne gagne rien. Ce qui signifie :

$$\forall (h \leftarrow B) \in \mathcal{R} :\ B \subseteq X \Rightarrow h \in X$$

$\emptyset$ est fermé ssi $\mathcal{R}$ ne contient aucun fait.

Définition 4 (cohérent/cofermé)**.** Un ensemble X est dit cohérent ou cofermé pour $\mathcal{R}$ ssi $X \subseteq \mathsf{T}_{\mathcal{R}}(X)$

Autrement dit : on ne perd rien.

Définition 5 (fixe)**.** Un ensemble X est dit fixe pour $\mathcal{R}$ si il est fermé et cohérent (cofermé) pour $\mathcal{R}$, c'est à dire ssi $\mathsf{T}_{\mathcal{R}}(X) = X$

Considérons l'ensemble de rêgles suivant :

$$
\begin{array}{ccc}
a \leftarrow & b \leftarrow a, b & e \leftarrow e \\
 & c \leftarrow a, c & \\
 & d \leftarrow b, c &
\end{array}
$$

et maintenant examinons certains ensembles pour voir s'ils sont fermés et/ou cohérents :

X	$\mathsf{T}_{\mathcal{R}}(X)$	**fermé**	**cohérent**
$\emptyset$	$\{a\}$	F	T
$\{a\}$	$\{a\}$	T	T
$\{a, b\}$	$\{a, b\}$	T	T
$\{a, d\}$	$\{a\}$	T	F
$\{d\}$	$\{a\}$	F	F

quels sont les ensembles fixes (points fixes) de ce système de rêgles ? Ce sont les ensembles suivants : $\{a\}$, $\{a, e\}$, $\{a, b\}$, $\{a, b, e\}$, $\{a, c\}$, $\{a, c, e\}$, $\{a, b, c, d\}$, $\{a, b, c, d, e\}$.

De manière générale, on a toujours un treillis complet de points fixes. Pourquoi ? D'abord pourquoi a-t-on un point fixe ? On se rapelle que $\mathsf{T}_{\mathcal{R}}(\emptyset)$ est l'ensemble des têtes de faits. On notera que pour tout X fixe, $\mathsf{T}_{\mathcal{R}}(\emptyset) \subseteq \mathsf{T}_{\mathcal{R}}(X) \subseteq \mathcal{H}(\mathcal{R})$. Donc, puisque $\mathsf{T}_{\mathcal{R}}$ est croissante, il existe un ensemble X entre $\mathsf{T}_{\mathcal{R}}(\emptyset)$ et $\mathcal{H}(\mathcal{R})$ tel que $\mathsf{T}_{\mathcal{R}}(X) = X$. Bien sur, si $\mathcal{R}$ ne contient aucun fait, alors $\emptyset$ est un point fixe.

Si X_1 et X_2 sont deux points fixes (donc $\subseteq \mathcal{H}(\mathcal{R})$), alors, puisque $\mathsf{T}_{\mathcal{R}}$ est croissante, $\mathsf{T}_{\mathcal{R}}(X_1) \subseteq \mathsf{T}_{\mathcal{R}}(X_1 \cup X_2)$, pareil pour $\mathsf{T}_{\mathcal{R}}(X_2)$. Or $\mathsf{T}_{\mathcal{R}}$ est croissante et bornée par $\mathcal{H}(\mathcal{R})$, donc il existe un ensemble Y entre $X_1 \cup X_2$ et $\mathcal{H}(\mathcal{R})$ tel que $\mathsf{T}_{\mathcal{R}}(Y) = Y$. On peut prouver l'unicité : c'est à dire qu'il existe un plus petit point fixe plus grand que X_1 et X_2.

On peut procéder de même vers le bas.

1.4 Ensemble défini inductivement

Il existe au moins un fermé pour $\mathcal{R}$, à savoir $\mathcal{H}(\mathcal{R})$. Donc on peut considérer l'intersection de tous les fermés, qu'on appelle $\mathrm{ind}(\mathcal{R})$.

Théorème 2 (ind($\mathcal{R}$) est fermé)**.** *ind(R), l'intersection de tous les fermés pour $\mathcal{R}$, est fermé pour $\mathcal{R}$*

Preuve. Soit X fermé pour $\mathcal{R}$. Alors, par définition de ind($\mathcal{R}$) comme le plus petit fermé, $\text{ind}(\mathcal{R}) \subseteq X$. Puisque $\mathsf{T}_{\mathcal{R}}$ est croissante, $\mathsf{T}_{\mathcal{R}}(\text{ind}(\mathcal{R})) \subseteq \mathsf{T}_{\mathcal{R}}(X)$. Donc $\mathsf{T}_{\mathcal{R}}(\text{ind}(\mathcal{R})) \subseteq \mathsf{T}_{\mathcal{R}}(X) \subseteq X$ car X est fermé. Puisqu'on a $\mathsf{T}_{\mathcal{R}}(\text{ind}(\mathcal{R})) \subseteq X$ pour tout X fermé :

$$\mathsf{T}_{\mathcal{R}}(\text{ind}(\mathcal{R})) \subseteq \cap\{X \mid X \text{ est fermé pour } \mathcal{R}\ \} = \text{ind}(\mathcal{R})$$

□

ind($\mathcal{R}$) est le plus petit (au sens de $\subseteq$) ensemble fermé pour $\mathcal{R}$.

Définition 6 (définition inductive)**.** ind($\mathcal{R}$) est l'ensemble défini inductivement par $\mathcal{R}$

On dit aussi que $\mathcal{R}$ est une définition par induction de cet ensemble. Dans l'example précédent $\text{ind}(\mathcal{R}) = \{a\}$. On notera que $\mathsf{T}_{\mathcal{R}}(\emptyset) \subseteq \text{ind}(\mathcal{R}) \subseteq \mathcal{H}(\mathcal{R})$, donc $\text{ind}(\mathcal{R}) = \emptyset$ ssi $\mathcal{R}$ ne contient aucun fait.

Nous verrons par la suite qu'on peut également définir ind($\mathcal{R}$) comme la limite d'une suite croissante d'ensembles : $X_0 = \emptyset$, $X_{i+1} = \mathsf{T}_{\mathcal{R}}(X_i)\ \forall i \in \mathbb{N}$.

1.5 Fermeture

La notion de fermeture, cloture, ou stabilité, est une variante de présentation.

Définition 7 (fermeture)**.** Soit $\mathcal{R}$ un ensemble de rêgles, et X un ensemble, la cloture $\text{Clo}_{\mathcal{R}}(X)$ d'un ensemble X par un système de rêgles $\mathcal{R}$ est le plus petit ensemble fermé pour $\mathcal{R}$ qui contient X.

Théorème 3. *C'est aussi l'ensemble défini inductivement en ajoutant à $\mathcal{R}$ tous les faits $x \leftarrow \emptyset$ pour tout $x \in X$:*

$$\text{Clo}_{\mathcal{R}}(X) = \text{ind}(\mathcal{R} \cup \{x \leftarrow \emptyset \mid x \in X\})$$

Inversement, un ensemble défini inductivement est la fermeture de $\emptyset$.

1.6 Démonstration par induction

C'est un procédé qui permet parfois de prouver que $\text{ind}(\mathcal{R}) \subseteq P$, c'est à dire que tous les éléments de ind($\mathcal{R}$) ont la propriété P.

- technique : on montre que P est fermé pour $\mathcal{R}$. Ceci prouve P pour les faits et propage cet invariant par les rêgles.
- suffisant : puisqu'on montre $\text{ind}(\mathcal{R}) \subseteq P$ et que $\text{ind}(\mathcal{R})$ est le plus petit fermé.
- méthode : Montrer que $\mathsf{T}_{\mathcal{R}}(P) \subseteq P$. Autrement dit : pour chaque rêgle $(h \leftarrow B) \in \mathcal{R}$, on montre que $B \subseteq P \Rightarrow h \in P$. En particulier, pour chaque fait, $h \in P$
- il se peut que P soit trop faible

La condition $B \subseteq P$ est apelée "hypothèse d'induction". Il se peut que $\text{ind}(\mathcal{R}) \subseteq P$ sans qu'on puisse le démontrer par cette méthode parce que P est trop faible : autrement dit, $\text{ind}(\mathcal{R})$ a une structure riche, mais l'invariant P ne reflète qu'un aspect assez pauvre de cette structure ; un aspect qui n'est pas suffisant pour prouver ce qui nous interesse vraiment. Il est donc souvent commode de prouver une propriété plus forte que P, c'est à dire avec plus de conditions restrictives.

Le remède, c'est de trouver un $F \subseteq P$ tel que F est fermé. S'il est vrai que $\text{ind}(\mathcal{R}) \subseteq P$, alors un tel F existe.

La méthode de démonstration par induction s'applique naturellement à la notion de fermeture en ajoutant les faits convenables.

1.7 Plus petit ensemble fixe

Théorème 4 (plus petit fixe)**.** $\text{ind}(\mathcal{R})$ *n'est pas seulement le plus petit fermé pour* $\mathcal{R}$*, c'est aussi le plus petit ensemble fixe pour* $\mathcal{R}$*.*

Preuve. Par définition, $\text{ind}(\mathcal{R})$ est fermé. On vérifie par induction que $\text{ind}(\mathcal{R})$ est cohérent pour $\mathcal{R}$. $\mathsf{T}_{\mathcal{R}}(\text{ind}(R))$ est fermé (prouvez le), c'est à dire :

$\mathsf{T}_{\mathcal{R}}(\mathsf{T}_{\mathcal{R}}(\text{ind}(\mathcal{R}))) \subseteq \mathsf{T}_{\mathcal{R}}(\text{ind}(\mathcal{R}))$. Or $\text{ind}(\mathcal{R})$ est fermé, donc $\mathsf{T}_{\mathcal{R}}(\text{ind}(\mathcal{R})) \subseteq \text{ind}(\mathcal{R})$, et puisque $\mathsf{T}_{\mathcal{R}}$ est croissante, $\mathsf{T}_{\mathcal{R}}(\mathsf{T}_{\mathcal{R}}(\text{ind}(\mathcal{R}))) \subseteq \mathsf{T}_{\mathcal{R}}(\text{ind}(\mathcal{R}))$ $\square$

Note : La fermeture de X par $\mathcal{R}$ n'est pas nécessairement fixe ; c'est seulement en ajoutant les $x \leftarrow \emptyset$ pour $x \in X$ qu'elle le devient.

On se referera au chapitre suivant pour d'autres informations concernant la notion de fermeture.

1.8 Examples de définitions inductives

Considérons la grammaire hors-contexte définie par les productions suivantes :

$$S \to \epsilon$$
$$S \to a\ S\ b$$

On a bien l'intuition qu'elle engendre le language $a^n b^n$: on le prouvera par induction. Pour cela, on se ramène à un ensemble infini de rêgles :

$$\begin{aligned} \epsilon &\leftarrow \emptyset \\ a\ m\ b &\leftarrow \{m\} && \forall m \in \{a, b\}^* \end{aligned}$$

Le language engendré par la grammaire est l'ensemble défini inductivement par ces rêgles. On montre par induction que c'est ensemble des mots de la forme $a^n b^n$ pour $n \in \mathbb{N}$.

1.9 Induction structurelle

Considérons la définition inductive des listes d'entiers dans un univers U de termes :

$$\begin{aligned} \mathsf{nil} &\leftarrow \emptyset \\ n{::}\ell &\leftarrow \ell && \forall n \in \mathbb{N},\ \forall \ell \in U \end{aligned}$$

Pour montrer, par induction structurelle, que toutes les listes d'entiers satisfont une propriété P, on montre $\mathsf{nil} \in P$ et $\ell \in P \Rightarrow n{::}\ell \in P$ pour tout $n \in \mathbb{N}$.

L'induction structurelle, c'est l'induction sur les rêgles qui caractérisent la formation des éléments d'un type de données.

1.10 Comparaison avec la récurrence

Récurrence simple. méthode de démonstration :

$$[0 \in P\ \wedge\ \forall n \in \mathbb{N}\,(n \in P \Rightarrow n + 1 \in P)] \Rightarrow \mathbb{N} \subseteq P$$

Preuve de correction : considérons $X = \{k \in \mathbb{N} \mid k \notin P\}$. Si X est non-vide, il a un plus petit élément, etc...

Cette méthode coincide avec la démonstration par induction sur $\mathcal{R}$ défini comme suit :

$$\begin{aligned} 0 &\leftarrow \emptyset \\ n + 1 &\leftarrow n && \forall n \in \mathbb{N} \end{aligned}$$

pourquoi ? parce que $\mathrm{ind}(\mathcal{R}) = \mathbb{N}$. preuve par induction.

Récurrence complète. (induction forte ur $\mathbb{N}$). méthode :

$$[0 \in P \ \wedge \ \forall n : (\forall m < n : \ m \in P) \Rightarrow n \in P] \Rightarrow \mathbb{N} \subseteq P$$

Cette méthode coindice avec l'induction sur $\mathcal{R}$ défini comme suit :

$$n \leftarrow \{m \in \mathbb{N} \mid m < n\} \qquad\qquad \forall n \in \mathbb{N}$$

C'est à dire l'ensemble infini des rêgles :

$$\begin{aligned} &0 \leftarrow \emptyset \\ &1 \leftarrow \{0\} \\ &2 \leftarrow \{0,1\} \\ &3 \leftarrow \{0,1,2\} \\ &\ \vdots \end{aligned}$$

Certaines démonstration se font plus naturellement par récurrence complète, par example : *tout entier* ≥ 2 *se décompose en produits de facteurs premiers.*

Attention ! La récurrence simple est un cas particulier de l'induction structurelle (utilise une construction particulière des entiers : celle de Peano).

La récurrence complète est un cas particulier de l'induction Noethérienne (utilise l'ordre sur $\mathbb{N}$).

1.11 Exercices

Exercice 1. *On considère la fonction factorielle* $f : \mathbb{N} \to \mathbb{N}$ *:*

$$f(x) = \textbf{\textit{si}} \ \ x = 0 \ \ \textbf{\textit{alors}} \ \ 1 \ \ \textbf{\textit{sinon}} \ \ x * f(x-1)$$

On peut la modéliser par l'ensemble infini de rêgles suivant :

$$\begin{aligned} (0,1) &\leftarrow \emptyset \\ (x+1, (x+1)*z) &\leftarrow \{(x,z)\} \qquad\qquad \forall x, z \in \mathbb{N} \end{aligned}$$

L'ensemble défini inductivement par ces rêgles est la fonction factorielle (on se rappelle qu'une fonction est une relation, c'est à dire un ensemble de couples $(x, f(x))$*).*

On montre par induction que cela calcule bien la factorielle ; c'est à dire que les couples sont de la forme $(n, n!)$*.*

Exercice 2. *On considère la définition d'une fonction* $f : (\mathbb{N} \times \mathbb{N}) \to \mathbb{N}$ *suivante :*

$$f(x, y) = \textbf{\textit{si}} \;\; x = y \;\; \textbf{\textit{alors}} \;\; 0 \;\; \textbf{\textit{sinon}} \;\; 1 + f(x + 2, y + 1)$$

Donner les règles de la définition inductive correspondante. En calculant quelques valeurs de f*, formuler une hypothèse de ce que* f *calcule. Prouver cette hypothèse par induction sur les règles de la définition inductive.*