

# A Fully Connectionist Model Generator for Covered First Order Logic Programs

Sebastien Bader, Pascal Hitzler Steffen Hölldobler, Andreas Witzel

## Contexte Général

L'étude interconnecte une représentation des connaissances symbolique (une partie de la logique du premier ordre) à une approche numérique (les réseaux de neurones). Le sous ensemble de la logique du premier ordre considéré est le programme couvert. Une clause couverte est une clause de Hornes dont la variable apparaissant dans le corps de la clause apparait aussi en tête.

## Motivation

Le raisonnement symbolique se base sur la connaissance de l'expert qui met à disposition son savoir dans les algorithmes. Un modèle satisfaisant le programme logique est trouvé par raffinement successif. Les réseaux de neurones encodent par leurs poids un autre une connaissance issu de l'apprentissage. Il est donc intéressant d'interconnecter ces deux approches disjointes.

## Travaux connexes

Ces mêmes auteurs se sont intéressés auparavant à transformer une interprétation logique en réel pour les réseaux de neurones. On peut citer les travaux : "Approximating the semantic of a Logic program" ou "Integrating First Order Logic Programs and connectionist system"

## Contribution

Une nouvelle représentation des interprétations en vecteurs de réels

Chaque symbole de prédicat décrit un axe. Les instances d'un symbole de prédicats sont ordonnées suivant la fonction de niveau et de dimension  $\iota$  :  $Interprétations \rightarrow Vecteurs\ de\ réels$ .  $\iota$  est bijective. L'étude considère ensemble les vecteurs issu d'axes différents mais qui ont le même niveau : les hyper cubes.

Un réseau de neurones qui approxime l'opérateur sémantique de la logique du 1er ordre

Soit  $Tp$  l'opérateur qui capture la sémantique d'un programme logique. Soit  $x$  un vecteur représentant un ensemble d'interprétation.  $Tp$  est intégré dans le réseau de neurones par la fonction  $fP(x) := \iota(Tp(\iota^{-1}(x)))$ .  $fP(x)$  est approximé par  $fQ(x)$  qui représente les interprétations en dessous d'un certain niveau.

Construction du réseau de neurones

Un neurone de la couche caché correspond à un Hyper Cube. Son vecteurs d'entrée correspond au centre de l'hyper cube. Son vecteur de sortie correspond à l'opérateur sémantique  $fQ$ . Les neurones d'entrée servent aux vecteurs  $x_{in}$  traduisant une interprétation. Les neurones de sortie capture la transformation  $fQ$

opéré par le réseau dans  $x_{out}$ , tel que  $x_{out}$  se rapproche du point fixe de  $Tp$  qui est un Modèle pour le programme.

## Une nouvelle méthode d'apprentissage

Lors de l'entraînement les poids en entrées et en sortie d'une unité de la couche sont ajustés en fonction de certains paramètres

$$wout \leftarrow \eta \cdot y + (1 - \eta) \cdot wout$$

$$win \leftarrow \mu \cdot c + (1 - \mu).$$

Une unité peut être retirée si elle ne s'active que très rarement. Une unité est ajoutée pour une région qui contient beaucoup d'erreurs

## Expérimentation

L'étude compare son approche à une autre méthode SVGN qui se base aussi sur la transformation des interprétations logique vers des réseaux de neurones. Les tests montrent pour les deux approches augmentent significativement le nombre d'unités dans la couche cachée.

Une comparaison est aussi faite entre une représentation par des vecteurs de 1 dimension et des vecteurs de dimension 2. Pour les vecteurs de dimensions 2 les tests mettent en relief une diminution du nombre d'unités liée à une augmentation du nombre d'erreurs lorsque le nombre d'exemples augmentent. L'opposé est observé pour les vecteurs de dimension 1.

## Résultats

Une fois le réseau construit et entraîné, les sorties sont liées aux entrées. Au bout de 6 itérations au maximum et à partir de n'importe qu'elle instances le réseau se stabilise sur un vecteur représentant le modèle.

## Conclusion et perspectives

L'étude présente une approche constructive qui approxime les opérateurs sémantiques de la logique du 1er ordre par la recherche d'un point fixe dans le réseau de neurones, toutefois les performances du réseau sont très dépendantes du choix des paramètres.

## Discussion :

### A propos des expérimentations

La comparaison avec les deux instances de leurs approches l'une utilisant les vecteurs de dimension 1 et l'autre des vecteurs de dimension 2, est difficilement interprétable puisque la paramétrisation s'est faite différemment. Il en va de même pour la comparaison de leur approche avec SVGN.

### A propos de l'approche

Une unité du réseau est responsable d'une région parallèle avec réseau connexionniste de reconnaissances des formes. La métaphore est bien trouvée.

Toutefois j'ai l'impression que la représentation adoptée par l'approche est reste très spécifique aux problèmes des programmes Logiques couverts voir à la représentation des nombres pairs dans cette logique.

L'article est difficile et pas assez clair sur certains points comme le rôle de l'opérateur sémantique.