

**AFPP – CINQUIÈME CONFÉRENCE INTERNATIONALE
SUR LES MÉTHODES ALTERNATIVES DE PROTECTION DES PLANTES
LILLE – 11 AU 13 MARS 2015**

OPTIMISER LE BIOCONTROLE DE LA CANNE À SUCRE PAR L'INFORMATISATION DE L'EXPERTISE

P. MARTIN^(1, 2), S. AUZOUX⁽¹⁾ et F.-R. GOEBEL⁽¹⁾

⁽¹⁾ CIRAD, UPR AïDA, F-34398 Montpellier Cedex 5 (France), pierre.martin@cirad.fr

⁽²⁾ LIRMM (CNRS – Université de Montpellier), F-34095 Montpellier Cedex 5 (France)

RÉSUMÉ

Comme d'autres filières agricoles, la filière canne à sucre cherche à utiliser des voies de protection alternative contre ses bioagresseurs, afin de produire mieux en regard du contexte économique, environnemental et sociétal d'aujourd'hui. L'objectif global est de réduire l'emploi des pesticides et de favoriser les solutions agro-écologiques, en particulier le biocontrôle. L'identification d'une méthode de lutte intégrée demande de concilier des connaissances concernant les pratiques culturales, le paysage, les bioagresseurs, les ennemis naturels et leurs interactions dans un environnement intensifié et simplifié (grands planteurs) ou composite par la présence d'une mosaïque de cultures et de végétation naturelle (petits planteurs). Pour résoudre cette problématique complexe et disposer d'un support de raisonnement, les connaissances expertes redevables de ce système ont été décrites sous la forme d'un réseau sémantique. Cet article présente la méthode utilisée et discute les premiers résultats.

Mots-clés : Protection intégrée des cultures, foreur de la canne à sucre, régulation naturelle, pratiques culturales, réseau sémantique, capitalisation des savoirs.

ABSTRACT

OPTIMIZING THE SUGAR CANE BIOCONTROL USING THE EXPERTISE COMPUTARIZATION

Like other agricultural sectors, the sugarcane industry seeks to use alternative ways of protection against pests to produce in a better way in line with the economic, environmental and societal context of today. The global aim is to reduce the use of pesticides and promote agro-ecological solutions, particularly biocontrol. Identifying an IPM strategy request a balance of knowledge about farming practices, landscape, pests, natural enemies and their interactions in a simplified environment (large-scale or commercial farms) or composite (small-scale farmers) by the presence of a mosaic of crops and natural vegetation. To solve this complex problem and have a reasoning support, expert knowledge of this system have been described as a semantic network. This paper presents the method and discusses the preliminary results.

Keywords: Crop protection, sugarcane borers, natural control, agricultural practices, semantic network, knowledge management.

INTRODUCTION

La canne à sucre est une culture tropicale à distribution mondiale cultivée dans plus de cent vingt (120) pays. Elle assure le revenu de nombreux planteurs ; et ses multiples produits (sucre, éthanol, bagasse, bioplastique, etc.) en font une culture économiquement importante pour les industries agroalimentaires et énergétiques (Botha, 2010). La production durable de canne et de sucre s'inscrit dans un contexte de réduction des intrants, en particulier de pesticides, ceci afin de produire mieux dans le contexte économique, environnemental et sociétal d'aujourd'hui. L'utilisation des voies de protection alternatives contre les bioagresseurs de la canne à sucre est ainsi de plus en plus privilégiée, comme les solutions agro-écologiques et en particulier le biocontrôle. Cette solution repose sur la gestion des équilibres des populations d'agresseurs, plutôt que sur leur éradication, au moyen de méthodes naturelles de protection des végétaux fondées sur les mécanismes et interactions qui régissent les relations entre espèces (Flint et Dreistadt, 1998). Différentes méthodes existent, à l'exemple de la lutte par conservation (Gonthier *et al*, 2014) ou les lâchers d'auxiliaires (Goebel *et al*, 2010).

La canne à sucre est attaquée par une grande diversité d'insectes, soit plus de mille cinq cent (1500) espèces dans le monde, à laquelle s'ajoutent plus de quatre vingt (80) maladies émanant de virus, champignons, bactéries, nématodes, phytoplasmes, etc. (Goebel et Sallam, 2011). Les insectes bioagresseurs dominant sont les vers blancs, les chenilles défoliatrices, les lépidoptères-foreurs et les piqueurs-suceurs (pucerons, cochenilles et thrips). Les lépidoptères foreurs sont à l'origine d'importants dégâts en Asie, en Afrique et en Amérique du Sud. Pour lutter contre ce type de ravageur, qui pénètrent dans la tige et se développent à l'intérieur, les insecticides sont proscrits car ils sont inactifs. On privilégie alors la résistance variétale et le biocontrôle, tel le lâcher inondatif de trichogrammes, un ennemi naturel minuscule qui se comporte en parasitoïde (Goebel *et al*, 2010).

L'identification d'une méthode de protection intégrée demande de concilier des connaissances sur les pratiques culturales, le paysage, les bioagresseurs, les ennemis naturels et leurs interactions dans un environnement simplifié (grands planteurs) ou composite (petits planteurs) par la présence d'une mosaïque de cultures et de végétation naturelle. Confronté à la diversité de techniques de biocontrôle disponibles et la pluralité de connaissances à maîtriser selon la zone cannière considérée, la problématique de l'agriculteur réside dans l'identification des bioagresseurs et les stratégies de protection (e.g. le biocontrôle) à mobiliser pour gérer au mieux la situation phytosanitaire locale dans son exploitation.

Actuellement, cette identification est effectuée par des experts dont le savoir est enrichi au gré des expériences dans les principaux pays producteurs de canne à sucre. Bien qu'interdisciplinaire par construction puisqu'il traite d'interactions entre organismes, le savoir de chaque expert est en majeure partie inféodé à un domaine disciplinaire privilégié (entomologie, pathologie, virologie, etc.). Ce faisant, confronté à une situation complexe où s'enchevêtrent des domaines disciplinaires distincts, dépassant de fait son champs de compétence, la réponse de l'expert reste partielle. Pour dépasser cette limite, la solution réside dans la composition des savoirs. La question traitée dans ce document est celle de l'élaboration d'une méthodologie de capitalisation de savoirs pluridisciplinaires permettant de produire un support de raisonnement pour l'expert.

La technique actuellement utilisée par la recherche pour capitaliser des connaissances repose sur la construction d'un modèle informatisé. Parmi les techniques de représentation des connaissances développées par la discipline informatique, la méthode des réseaux sémantiques s'intéresse à formaliser les connaissances exprimées selon le langage naturel sous la forme d'un graphe (Sowa, 1984). De nombreux auteurs ont utilisé cette méthode pour décrire un système observé, par une ou plusieurs disciplines, sous forme d'ontologie, e.g. respectivement Shrestha *et al* (2012) et Athanasiadis *et al* (2009). D'autres ont utilisé ce formalisme pour produire des graphes de

connaissance, à l'exemple d'Agropedia (2014) qui a modélisé un grand nombre de cultures agricoles présentes en Inde. Bien que ces graphes comportent des connaissances relatives au biocontrôle, aucune logique de traitement automatique du modèle n'est utilisée. Ce faisant, le graphe-modèle est un dessin comportant des chemins que l'utilisateur doit parcourir, et composer lui-même, pour élaborer son raisonnement.

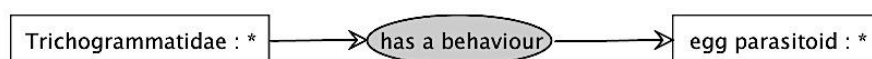
Dans l'objectif de proposer un modèle interactif, qui permet à l'auteur de tester des hypothèses sur son système, nous avons choisi d'utiliser un autre type de réseau sémantique, celui des Graphes Conceptuels (Chein et Mugnier, 2009). Cette méthode, qui est une alternative aux Logiques de Description (Baader *et al*, 2003), met à disposition des développeurs de modèles le logiciel Cogui (Gutierrez, 2014). Celui-ci est simple d'usage et permet d'effectuer des raisonnements, à l'exemple du test d'hypothèses, sur les graphes de connaissance. L'objet de cet article est de présenter la méthode utilisée pour représenter les connaissances d'un entomologiste Expert dans le biocontrôle de la canne à sucre (et plus largement dans la protection intégrée), et les premiers résultats.

MATERIEL ET METHODE

Représenter Un Graphe De Connaissance Et Reasonner Avec Les Graphes Conceptuels

La méthode des Graphes Conceptuels propose un formalisme de représentation et de raisonnement sur les connaissances élaboré selon la théorie des graphes. La représentation des connaissances s'effectue sous la forme d'un graphe orienté dans lequel un nœud représente un concept et un arc reliant des nœuds une relation sémantique entre les concepts reliés. La figure 1 présente un exemple de représentation de connaissances. Celle-ci stipule qu'un Trichogramme a le comportement d'un parasitoïde d'œuf.

Figure 1 : Exemple de représentation de connaissances. Dans cette figure, un rectangle représente un concept et un ovale la relation sémantique entre les concepts
Example of knowledge representation. In this figure, a rectangle represents a concept and an oval the semantic relation between the concepts

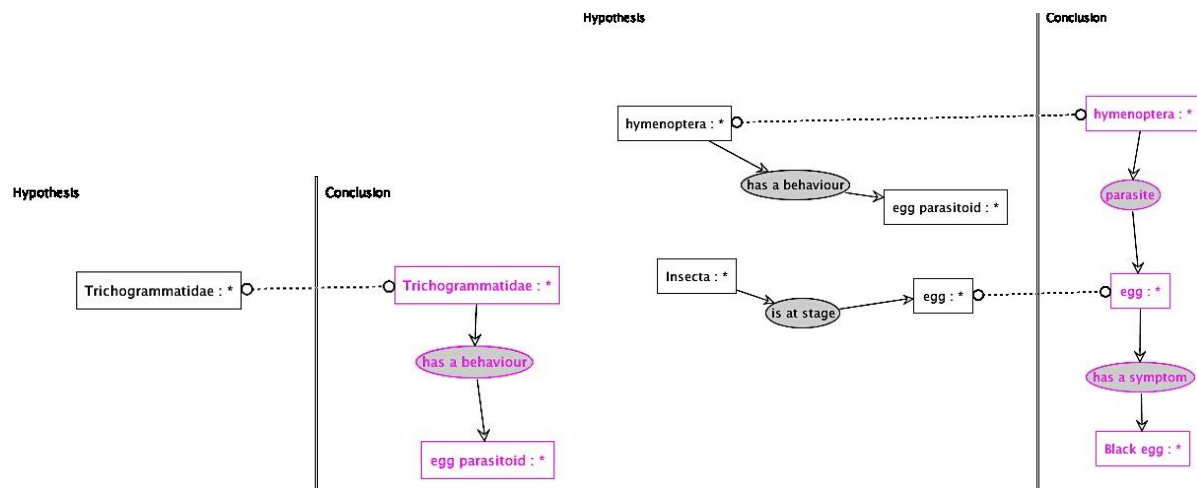


En préalable à la production de graphes de connaissance, ce formalisme nécessite la construction de deux vocabulaires, un pour les concepts et un autre pour les relations sémantiques. Ces vocabulaires comportent l'ensemble des éléments retenus pour exprimer les connaissances. Au sein de ces vocabulaires, les éléments, i.e. concepts ou relations sémantiques, sont reliés entre eux au moyen de la relation de généralisation (« is a »). Du point de vue sémantique, cette relation permet de préciser la signification d'un élément. Un exemple de généralisation pourrait être qu'un trichogramme est un hyménoptère. Cet exemple serait implémenté au moyen de deux concepts « Trichogrammatidae » et « Hymenoptera » reliés par une relation « is a » orientée du premier vers le second. Par construction, de part l'application successive de la relation de généralisation entre les éléments, les deux vocabulaires peuvent présenter une structure arborescente. Dans ce cas, le concept situé en amont du vocabulaire porte la sémantique la plus générale (ou englobant) du vocabulaire.

Dans notre travail, deux types de graphes de connaissance proposés par les Graphes Conceptuels ont été utilisés. La Règle d'Inférence permet de décrire les connaissances selon la forme déductive « Si ...

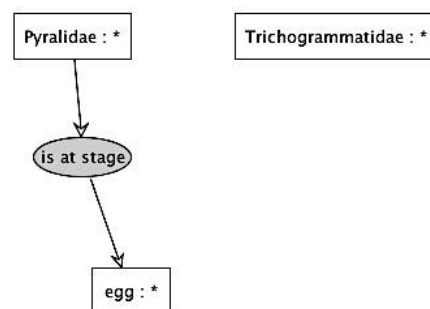
Alors... ». La Figure 2 présente deux exemples de Règle d'Inférence. La règle de gauche stipule que s'il y a un trichogramme (hypothèse) alors celui-ci a un comportement de Parasitoïde d'œuf (conclusion). La règle de droite décrit l'effet d'un hyménoptère ayant le comportement de parasite d'œuf mis en présence d'un insecte au stade œuf : le premier parasite le second, ce dernier présentant alors le symptôme appelé « black egg ».

Figure 2 : Deux exemples de Règle d'Inférence
Two examples of Inference Rule



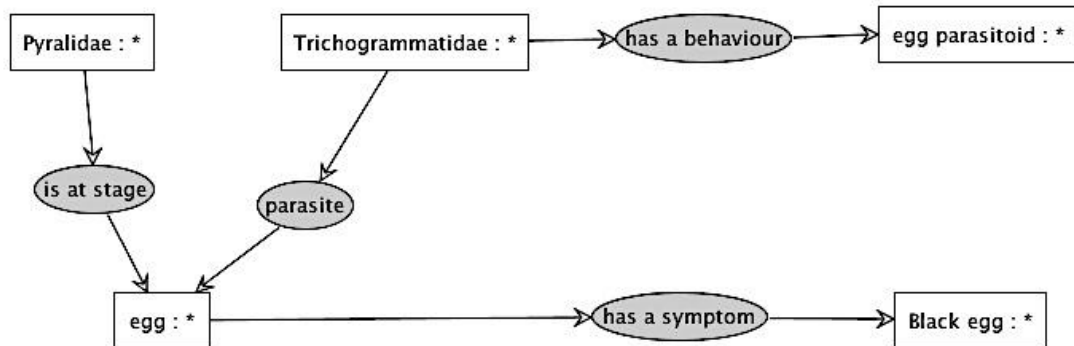
L'autre graphe de connaissance est le Graphe de Fait. Celui-ci permet, dans le cadre de ce travail, de spécifier les éléments constitutifs du paysage étudié. La Figure 3 présente un exemple de paysage simplifié constitué d'un Pyralidé au stade œuf et d'un Trichogramme.

Figure 3 : Exemple de Graphe de Fait
Example of Fact Graph



In fine, le raisonnement sur les graphes de connaissance consiste à appliquer les Règles d'Inférence sur le Graphe de Fait pour produire un Graphe de Raisonnement. Cette opération de raisonnement est réalisée de façon automatique par le logiciel Cogui. La Figure 4 présente l'application des Règles d'Inférence présentées en Figure 2 sur le Graphe de Fait de la Figure 3.

Figure 4 : Graphe de Raisonnement résultant de l'application des Règles d'Inférence de Figure 2 sur le Graphe de Fait de la Figure 3
Reasoning Graph resulting from the application of the Inference Rules presented in Figure 2 on the Fact Graph presented in Figure 3



Méthodologie D'extraction Et De Formalisation Des Connaissances

Pour produire cette activité de modélisation, le groupe de travail mis en place était constitué de trois membres : l'Expert, un modélisateur spécialiste des réseaux sémantiques, et un informaticien ayant une compréhension partielle du système cannier et des réseaux sémantiques. Le recueil et la transcription des connaissances ont été réalisés conjointement dans le cadre de séances de travail d'une durée de deux heures.

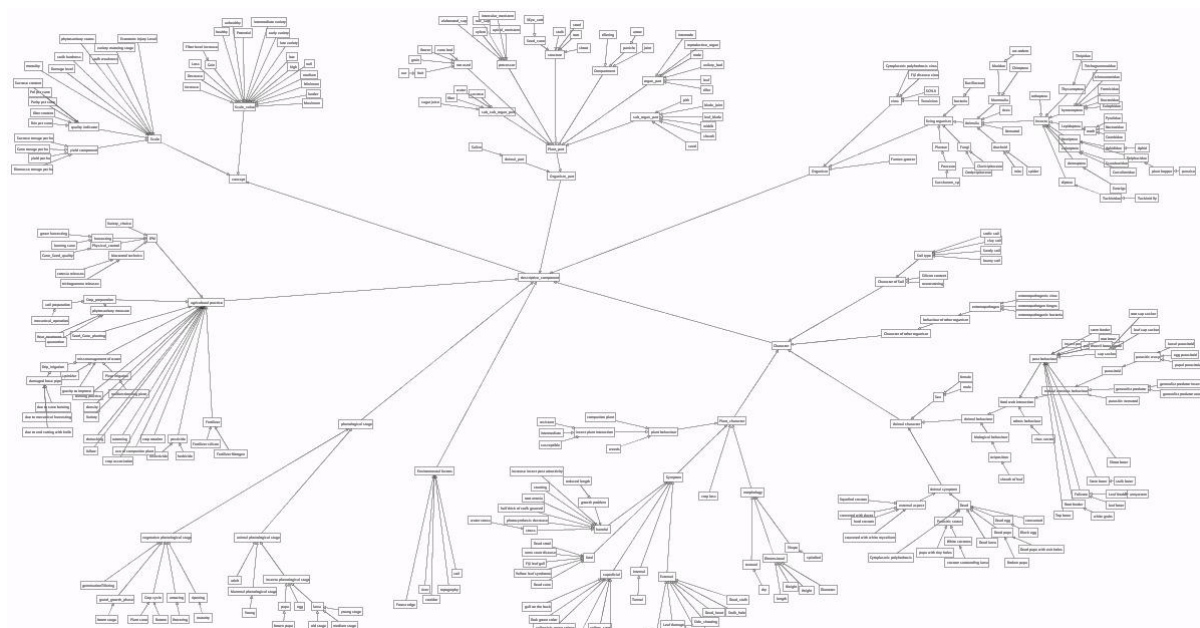
Les séances devaient permettre de recenser les mécanismes et interactions entre organismes ayant un impact sur la santé de la canne à sucre. Les règles de fonctionnement du groupe étaient les suivantes :

- Les considérations agronomiques n'ayant pas de lien avec la protection de la culture (e.g. effet de l'azote sur le rendement) n'ont pas été considérées.
- Pour structurer des parties de vocabulaire, la bibliographie a été utilisée autant que faire se peut
- Les connaissances ont été représentées en langue anglaise, dans l'objectif d'anticiper le transfert de technologie auprès des partenaires.

RESULTAT

Après seize (16) séances de travail, le vocabulaire des concepts comporte trois cent quarante et un (341) éléments (Figure 5). Ce graphe possède une structure arborescente. Le concept le plus général s'intitule « Concept ». Les 340 autres concepts sont répartis selon sept domaines distincts (cf. Tableau 1). Concernant le domaine des organismes, celui-ci comporte cinquante cinq (55) concepts réparties selon différents sous-domaines : Virus (5), plantae (3), fungi (3), bacteria (2), arachnid (3), mamalia (4), aves (1), nematod (1) et insectae (29).

Figure 5 : Le vocabulaire des concepts tel qu’affiché par Cogui. Dans cette figure, un rectangle représente un concept et une flèche une relation de généralisation
Vocabulary concepts as displayed by Cogui. In this figure, a rectangle and an arrow represent respectively a concept and generalization relationship



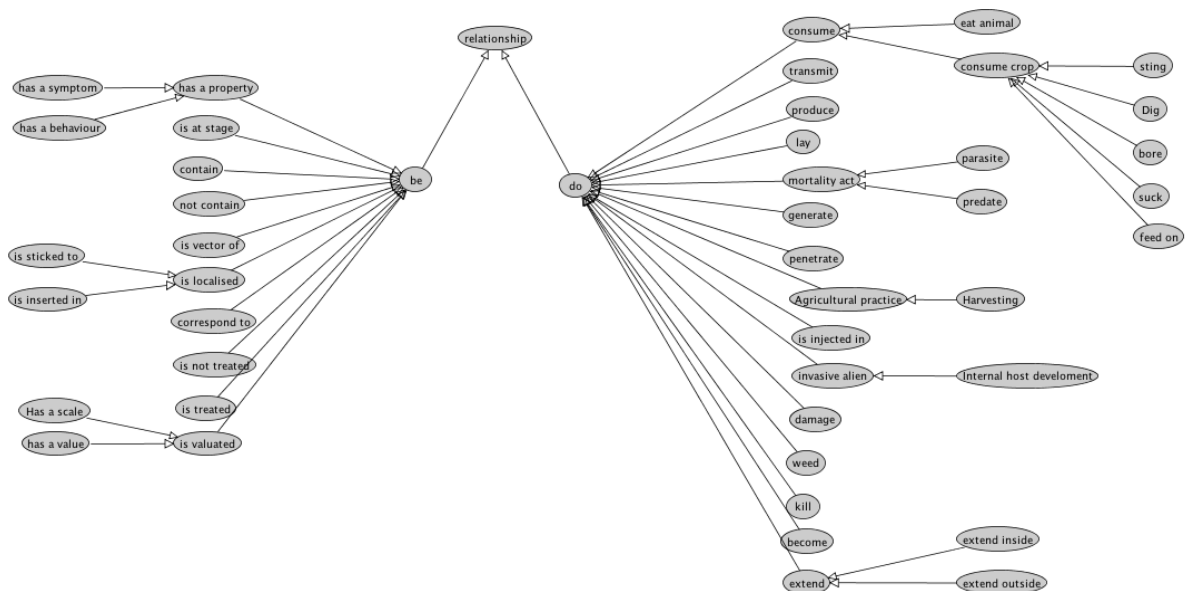
Deux éléments de bibliographie ont permis de structurer une partie de ce vocabulaire : Bourgoin (2014) a permis de structurer les organismes selon les liens taxonomiques, tels que définis par la communauté scientifique, quant à Moore & Botha (2013), il a servi à produire la description de la structure de la plante et ses stades de développement.

Tableau 1 : Les sept domaines du vocabulaire des concepts
The seven domains of the concept vocabulary

Domaine	Nb de concepts
Les organismes (insectes d’ordre divers, araignées, champignons, bactéries, mammifères, oiseaux, virus, etc.)	55
Les éléments constitutifs des organismes (tige, feuille, salive d’insecte, etc.)	46
Les stades phénologiques de la plante et de certains organismes (boom stage, ratoon, egg, pupa, etc.)	24
Les pratiques culturelles (méthodes de biocontrôle, préparation du sol, mauvaise gestion de l’eau, fertilisation, etc.)	42
Les facteurs environnementaux (lisière de forêt, rivière, corridor, topographie, sol)	6
Les caractéristiques des organismes (comportement, symptôme, etc.)	128
La quantification des mécanismes et des interactions au moyen de deux concepts, i.e. l’échelle (mortalité, indicateur de qualité, etc.) et la valeur sur l’échelle (high, medium, low, etc.).	39

Le vocabulaire des relations comporte quarante six éléments (cf. Figure 6). Ce vocabulaire possède également une structure arborescente. La relation sémantique la plus générale, située en amont de l'arborescence, s'intitule « relationship ». En accord avec les travaux de Tesnière (1986) sur le langage naturel, les relations sémantiques ont été distinguées selon deux groupes : celles permettant d'indiquer un état (« to be » en Figure 6) et celles descriptives d'une action (« to do » en Figure 6). Ces deux groupes comportent respectivement dix sept (17) et vingt huit (28) relations sémantiques. Parmi cet ensemble de relations sémantiques, on peut noter que, dans le premier groupe, deux relations i.e. « not contain » et « is not treated » sont précédées de l'adverbe « not » exprimant la négation.

Figure 6 : Le vocabulaire des relations sémantiques tel qu'affiché par Cogui. Dans cette figure, un ovale représente une relation sémantique et une flèche une relation de généralisation
The vocabulary of semantic relationships as displayed by Cogui. In this figure, an oval and an arrow represent respectively a semantic relation and a generalization relationship



Ces deux vocabulaires ont été enrichis au gré de l'implémentation des Règles d'Inférence. Leur construction a été réalisée par thématique pour donner lieu à des bibliothèques de Règles d'Inférence. La première thématique considérée concernait les relations trophiques, depuis les ennemis naturels vers la plante, et la seconde l'impact des pratiques agricoles sur les organismes. Leur implémentation a été effectuée sous la forme de cent huit (108) Règles d'Inférence, réparties au sein de cinq bibliothèques (cf. Tableau 2). La production de ces règles a nécessité l'utilisation de sept cent quarante huit (748) concepts et quatre cent quarante sept (447) relations sémantiques.

Tableau 2 : Les cinq bibliothèques de Règles d'Inférence
The five libraries of Inference Rules

Bibliothèque	Nb de règles
Croissance des insectes par stade physiologique	2
Modèle comportemental de chaque organisme (entomopathogène, piqueur-suceur, etc.)	30
Description physique des comportements	38
Croissance de la plante par stade physiologique	6
Impact des pratiques agricoles sur les organismes (irrigation par asperseur ou par pivot, fertilisation, récolte, implantation de la culture, préparation du sol, gestion de la canne, choix variétal, etc.)	32

Une bibliothèque supplémentaire, appelée Bibliothèque de Gestion, a été développée par le modélisateur au gré des besoins du groupe de travail. A titre d'exemple, parmi les règles produites, il y en a une qui retire les sous-graphes redondants au sein d'un graphe de connaissance, facilitant de fait la lecture du Graphe de Raisonnement.

DISCUSSION

Au terme de trente quatre heures (34) de travail, les séances ont permis de recenser et formaliser une part des connaissances de l'Expert. La répartition des concepts au sein du vocabulaire informe des thématiques couvertes par les connaissances formulées par l'Expert. En particulier, une majorité de concepts du domaine des organismes relèvent du sous-domaine « Insectae », soit vingt-neuf (29) concepts sur les cinquante cinq (55) présents. Les autres concepts étant distribués dans des sous-domaines très différents (fungi, bacteria, arachnid, etc.). Ce faisant, on peut en déduire que les connaissances implémentées font majoritairement référence à l'entomologie, les autres sous-domaines correspondant pour la plupart à des auxiliaires de ces insectes.

Par ailleurs, le vocabulaire des relations sémantiques comporte deux relations utilisant la négation. Bien que cette forme de représentation ne respecte pas la convention adoptée par la théorie des réseaux sémantiques, ce qui constitue donc à une anomalie de modélisation, cette formulation a été conservée puisqu'elle est conforme à la connaissance de l'Expert. Par contre, pour que le raisonnement puisse être réalisé, une Règle d'Inférence, transformant de façon automatique cette écriture interdite selon un format manipulable par la théorie des graphes, est implémentée dans la bibliothèque de gestion.

Pour être formalisées sous forme de Graphe Conceptuels, les connaissances ont été exprimées sous forme de Règles d'Inférence. Pour décrire les réseaux trophiques, la connaissance de l'Expert a été répartie dans deux bibliothèques : le comportement des organismes et la description physique des comportements. Ce faisant, la description d'une relation trophique, du type « x mange y sous certaines conditions », résulte de la composition de Règles d'Inférence. Cette forme de représentation des connaissances résulte de deux apprentissages :

- la compréhension des mécanismes biologiques par le modélisateur pour proposer un modèle de transcription des connaissances de l'Expert selon un formalisme qui souscrive à la logique mathématique sous-jacente aux Graphes Conceptuels,
- et la compréhension de la logique de fonctionnement des Graphes Conceptuels par l'Expert pour être en mesure de décomposer ses connaissances selon le formalisme proposé par le modélisateur.

Cet apprentissage, qui a nécessité deux séances de travail avant d'accéder à une compréhension mutuelle, a été facilité par l'informaticien. En effet, celui-ci a joué le rôle d'interprète pour les deux parties, voire demander à chacun, aux grès des échanges, d'explicitier l'implicite de ses propos.

CONCLUSION

Les connaissances de l'Expert implémentées jusqu'à présent concernent les relations trophiques entre les organismes ainsi que l'impact des pratiques agricoles sur ces organismes. Cet ensemble suffit pour appréhender les stratégies de biocontrôle du point de vue entomologique à l'échelle d'une parcelle cultivée. Pour que cette solution agro-écologique soit appréhendée dans son intégralité, la base de règles d'Inférence doit être étendue aux connaissances d'autres disciplines (malherbologie, pathologie, etc.) puis complétée des connaissances reliant l'environnement avec les organismes (nuisibles et auxiliaires).

L'étude du biocontrôle d'une culture de canne à sucre, pour une situation donnée, consistera alors à produire un Graphe de Fait comportant les organismes observés à différentes étapes d'évolution du paysage. Différentes méthodes d'optimisation du biocontrôle ont déjà été évoquées, à l'exemple de la production de Règles d'Inférences permettant d'ajouter, extraire et déplacer de façon automatique certains organismes ou cultures agricoles présents dans le paysage.

L'aboutissement de ce travail conduira à disposer d'un outil d'aide au raisonnement. Plusieurs utilisations sont envisagées par l'Expert, à l'exemple de support d'explication des solutions agro-écologiques proposées aux agriculteurs. La pérennité de ce travail est conditionnée par la mise à jour de la base de connaissances au gré de l'évolution des techniques culturales et des savoirs des experts. Cette opération sera aisée grâce à la simplicité du formalisme des Graphes Conceptuels et la souplesse d'utilisation de Cogui.

BIBLIOGRAPHIE

- Agropedia, 2014. Agropedia: an online knowledge repository for information related to agriculture in India. Ressource internet (<http://agropedia.iitk.ac.in/>), consultée le 5 novembre 2014.
- Athanasiadis I.N., Rizzoli A.-E., Janssen S., Andersen E., Villa F., 2009. *Ontology for Seamless Integration of Agricultural Data and Models*. 3rd Intl Conf on Metadata and Semantics Research (MTSR'09), (Sartori, F., Sicilia, M. A., and Manouselis, N., Eds.), SpringerVerlag, 282-293.
- Baader F., Calvanese D., McGuinness D. L., Nardi D., Patel-Schneider P. F., 2003. *The Description Logic Handbook: Theory, Implementation, Applications*. Cambridge University Press, Cambridge, UK, 624p.
- Botha F.C., 2010. Future prospects. *Genetics, Genomics and Breeding of Sugarcane*, 31, edited by Henri R, Kole C. Science Publishers, 249-264.
- Bourgoin T., Baillargeon G., Decock W., De Wever A., Didžiulis V., 2014. Species 2000 & ITIS Catalogue of Life, 20th January 2014. Ressource internet (<http://www.catalogueoflife.org/col>), Species 2000: Naturalis, Leiden, the Netherlands.
- Chein M., Mugnier M.-L., 2009. *Graph-based Knowledge Representation and Reasoning - Computational Foundations of Conceptual Graphs*. Advanced Information and Knowledge Processing Series, Springer, London, 427 p.
- Flint M. L., Dreistadt, S. H., 1998. *Natural Enemies Handbook: The Illustrated Guide to Biological Pest Control*. University of California Press, 154 p.
- Goebel F.R., Tabone E., Do Thi Khanh H., Roux E., Marquier M., Frandon J., 2010. Biocontrol of *Chilo sacchariphagus* (Lepidoptera: crambidae) a key pest of sugarcane: lessons from the past and future prospects. *Sugar Cane International*, 28, 128-132.

Goebel F.R., Sallam N., 2011. New pest threats for sugarcane in the new bioeconomy and how to manage them. *Current Opinion in Environmental Sustainability*, 3, 81-89.

Gonthier J.G., Ennis K.K., Farinas S., Hsieh H-Y, Iverson A.L., Batary P., Rudolphi J., Tschardt T., Cardinale B.J., Perfecto I., 2014. Biodiversity conservation in agriculture requires a multi-scale approach. *Proceedings of the Royal Society*, B 281, 20141358.

Gutierrez, 2014. *Cogui : a conceptual graph editor*. Ressource internet (<http://www.lirmm.fr/cogui/>), Consulté le 5 novembre 2014.

Moore P. H., Botha F. C., 2013. *Sugarcane: physiology, biochemistry & functional biology*. Wiley-Blackwell, 693 p.

Shrestha R., Matteis L., Skofic M., Portugal A., McLaren G., Hyman G., Arnaud E., 2012. Bridging the phenotypic and genetic data useful for integrated breeding through a data annotation using the Crop Ontology developed by the crop communities of practice. *Frontiers in Physiology*, 3, 0326.

Sowa J. F., 1984. *Conceptual Structures: Information Processing in Mind and Machine*. Addison-Wesley, 481 p.

Tesnière L., 1988. *Éléments de syntaxe structurale*. Ed. Klincksieck, Paris, 674 p.