

# NOTICE INDIVIDUELLE ET CURRICULUM VITÆ

---

NOM : BENZAKEN

PRÉNOM : VERONIQUE

DATE ET LIEU DE NAISSANCE : 16 juin 1960 à Tarascon (13)

NATIONALITÉ : Française

SITUATION DE FAMILLE : Célibataire, une fille (1999)

PROFESSION : Professeur

ADRESSE : LRI, Université de Paris-Sud, bat 490, 91405 Orsay Cedex

TÉLÉPHONE : (bureau) 01 6915 6628, (portable) 06 8392 0370

ADRESSE ÉLECTRONIQUE : veronique.benzaken@lri.fr

En bref:

## Situation Professionnelle.

**1998** Depuis septembre 1998, **Professeur** à l'Université de Paris Sud 11. **Promue à la première classe** depuis 2004, en CRCT au titre local pour 2003-2004. **Titulaire de la prime d'encadrement doctoral** d'octobre 1993 à novembre 2001 puis depuis octobre 2006. Membre de l'équipe Bases de Données du LRI (Laboratoire de Recherche en Informatique) UMR 8623 du CNRS.

Professeur (1998) Paris 11, Maître de conférences (1990) Paris 1.

**1990-1998** D'octobre 1990 à août 1998, **Maître de Conférences** à l'Université de Paris 1 (Panthéon-Sorbonne).

**1990-1998** De 1990 à 1998, Chercheur au CRI (Centre de Recherche en Informatique) de Paris 1.

**1990-1991** De 1990 à 1991, **Conseiller extérieur** à l'INRIA.

## Études et diplômes.

**1996** Habilitation à diriger des recherches - Université Paris 11.

JURY : C. Delobel Président, M. Adiba Rapporteur, A. Albano Rapporteur, M.P. Atkinson Rapporteur, M-C. Gaudel, F. Gire, C. Rolland, N. Spyrtos.

TITRE : "Langages de programmation pour bases de données et contraintes d'intégrité".

HDR (12/12/96) et Doctorat (16/1/90) Paris 11, DEA (15/9/86) Paris 7 ; Premier Prix de chant et d'art-lyrique, (15/6/83) CNR Grenoble.

**1990** Thèse de Doctorat, effectuée sous la direction de Mr le Professeur C. Delobel.

JURY : C. Delobel Président, D. J. DeWitt Rapporteur, P. Valduriez Rapporteur, F. Bancelhon, J. Beauquier.

TITRE : "Regroupement d'objets sur disque dans un système de bases de données orienté-objet". Spécialité Informatique, Université de Paris 11.

**1986** D.E.A. d'Informatique Fondamentale - Université Paris 7. Mention : Très bien.

**1983** Premier prix de Chant et Premier prix d'Art-Lyrique du Conservatoire National de Région (CNR) de Grenoble.

## Expérience industrielle.

**1984** Ingénieur d'application, Société ORDINA.

**1986-1989** Dans le cadre de ma thèse sous convention CIFRE, Ingénieur, Société IN2.

Ingénieur d'application SSII (1984) ; Convention CIFRE (86-90)

## Enseignement.

Je ne donne pas le détail de mes activités d'enseignement assurées en qualité de maître de conférences. J'ai enseigné dans tous les cycle de l'Université tant des cours théoriques (théorie des langages et compilation) que des cours plus pratiques (bases de données, programmation). Je décris mes activités effectuées en qualité de professeur.

### 1998 à ce jour : Professeur Université de Paris 11

#### Niveau L

**Licence d'Informatique et IUP Miage.** cours "Introduction aux bases de données". Université de Paris 11 ( 1998-2003).

**Licence Miage par apprentissage** cours "Bases de données". Université de Paris 11, (1998-2001).

**Licence Miage classique.** TD du cours "Bases de données". Université Paris 11, (1998-1999) et (2001-2002).

**C<sub>4</sub>-DU** Diplôme d'Université en Informatique, cours de "Bases de données", Université de Paris 11, (2001-2003).

**Licence de Bio Informatique** cours de "Bases de Données". Université de Paris 11, ( 2002-2003).

**IUP-Miage** Tutrice de dix apprentis et encadrement de stages annuellement.

**Miage : VAE** Responsable de trois professionnels dans le cadre de la VAE.

#### Niveau M : Professionnel

**Ecole d'ingénieurs IFIPS (niveau M2)** Cours aspects systèmes des bases de données : concurrence, reprise, distribution et tuning. Depuis 2008.

**Master Pro M2** Cours de Gestion de l'information sur Internet (12h) depuis 2004

**Ecole d'ingénieurs IFIPS (niveau M1)** Cours de Bases de données avancées (20h en apprentissage et 20h en formation classique). Depuis 2004. La page décrivant le contenu du cours est à <http://www.lri.fr/benzaken/ifips.en.html>

**Ecole d'ingénieurs IFIPS (niveau M2)** Cours de Nouvelles architectures logicielles (15h), de 2004 à 2008.

**Maîtrise Miage par apprentissage** cours "Objets, méthodes et langages" Université Paris 11, (1998-2000).

**Encadrements de stages** Depuis 1998 à raison de cinq à six stagiaires par an en moyenne.

**DESS Ingénierie Informatique** Université de Paris 11, cours "Bases de données avancées", (2000-2001).

**Dess** Encadrement de dix stages (année 2000-2001).

Enseignement dans plusieurs écoles doctorales en Italie; enseignement en DEA (Paris 1, Paris 11), DESS (Paris 11), cours et travaux dirigés en DEUG Mass, Licence et Maîtrise d'informatique, Miage, Mass et DU.

Niveau M : Recherche

**Master Recherche** Cours Langages et systèmes pour bases de données (depuis 2004).

**DEA d'Informatique I3: information, interaction, intelligence** Universités Paris 1, Paris 11. Cours "Sûreté, Intégrité et Sécurité des Données" en collaboration avec G. Castagna (1998-2003).

**DEA d'Informatique I3: information, interaction, intelligence** Universités Paris 1, Paris 11. Cours "Fondements des bases de données", en collaboration avec N. Bidoit et N. Spyrtatos(1999-2003)

**DEA Théorie et Ingénierie des Bases de Données.** Universités Paris 11 - Paris 1. Module "Langages de Programmation pour Bases de Données". 1998-1999.

Écoles doctorales

**Scuola Nazionale dei Dottorati di Informatica delle Facoltà di Scienze.** Cours "Linguaggi per basi di dati ad oggetti", Bertinoro (Italie), 1999 (8 heures) [en Italien].

**Responsabilités collectives.**Période 1990-1998 : Maître de conférences Université de Paris 1

1. Membre du **Conseil d'Administration** de l'Université de Paris 1 de mars 1993 à mars 1997.
2. Membre du **Conseil de l'U.F.R. de Mathématiques et Informatique** de Paris 1 de Juin 1992 à Septembre 1998.
3. Membre du **Bureau du Conseil de l'U.F.R. de Mathématiques et Informatique** de Paris 1 d' Octobre 1992 à Septembre 1998.
4. Responsable, pour l'U.F.R., de 1992 à 1994, de **l'organisation des enseignements d'Informatique** dans les diverses filières de Paris 1.
5. **Assesseur** de la **Commission de Spécialistes** Mathématiques et Informatique de l'Université de Paris 1, d' Avril 1995 à Septembre 1998.
6. Membre de la **Commission de Spécialistes** Mathématiques et Informatique de l'Université de Paris 1 de Septembre 1993 à Septembre 1998.
7. Membre de la **Commission de Spécialistes** Informatique de l'Université d'Orléans de Mai 1994 à Mai 1995.
8. Membre de la **Commission de Spécialistes** Informatique de l'Université de Paris 12 (Créteil), de Janvier 1996 à Septembre 1998.
9. Membre de la **Commission de Spécialistes** Informatique de l'Université de Versailles de Décembre 1996 à Décembre 1997.
10. Membre de la **Commission de documentation** de l'Université de Paris 1 d' Avril 1994 à mars 1997.

Période 1998 à ce jour : Professeur Université de Paris 11

1. Membre élue du **CTPU**, de décembre 2003 à 2006.

Maître de conférences :

Membre du conseil d'administration, membre de 4 commissions de spécialistes, membre du conseil d'UFR et de diverses commissions de l'Université.

Professeur :

Membre du conseil de département, bureau du département enseignement, membre du conseil de laboratoire du LRI, membre de la commission de spécialistes, directrice de l'IUP-Miage.

2. Membre titulaire de la **Commission de Spécialistes** Informatique de l'Université de Paris 11 depuis décembre 2003.
3. **Directrice de l'IUP-Miage** de 1998 à 2003.
4. Membre du **Conseil du département enseignement** du département d'Informatique de Paris 11 de 1998 à 2003.
5. Membre du **Bureau du Conseil du département enseignement** du département d'Informatique de Paris 11 de 1998 à 2003.
6. Membre du **Conseil du laboratoire** du LRI, UMR 8623 CNRS de 1999 à 2008..
7. Membre suppléant de la **Commission de Spécialistes** Informatique de l'Université de Paris 11 de mars 2001 à décembre 2003.
8. Membre du **Conseil Scientifique** de l'Université de Paris 11, de mars 2002 à mars 2005.
9. Membre du **Conseil de la recherche de l'UFR des Sciences** de mars 2004 à mars 2005..

*Globalement, j'ai consacré quatre jours par semaine en moyenne à la gestion pédagogique et administrative de l'IUP. Les autres responsabilités ne m'ont occupée qu'une demi journée par mois.*

## Recherche.

Mes thématiques de recherche sont à la confluence de deux domaines : les bases de données et les langages de programmation. L'objectif visé étant l'obtention tant de langages prenant en compte des volumes massifs de données que de systèmes ou plateformes offrant un environnement de programmation robuste, fiable et efficace. À cette fin, nous étudions, systématiquement, les fondements théoriques permettant d'atteindre de tels objectifs (en particulier analyse statique, interprétation abstraite, théorie des types, déduction automatique) sans jamais négliger les aspects relatifs à l'implantation de prototypes et/ou de produits robustes.

### Doctorat

#### Aspects système des systèmes de gestion de bases de données

En 1987, j'ai rejoint, dans le cadre de ma thèse, le GIP Altaïr (INRIA). J'ai participé à la conception puis à la réalisation du système de gestion de bases de données orientées objets O<sub>2</sub>. Ma contribution a particulièrement porté sur des problèmes de gestion du disque [30, 41, 40, 28, 39, 29, 3] et a consisté à définir et implanter des politiques de placement de données afin de minimiser les surcoûts d'entrée-sortie.

À partir de 1989, le GIP s'est scindé en deux équipes : l'une chargée de l'industrialisation du prototype O<sub>2</sub>, l'autre poursuivant une activité de recherche. Au sein de cette dernière équipe, j'ai commencé une activité de recherche plus fondamentale. Dans le cadre du projet ESPRIT II "Basic Research Action" n° 3070 FIDE<sup>1</sup>, démarré en 1989, j'ai étudié plus en détail les problèmes de la définition des systèmes de gestion d'objets persistants et des langages de programmation pour bases de données. Ce projet m'a permis de collaborer avec des équipes internationales de pointe sur le domaine des bases de données et des langages de programmation persistants (Equipe de Malcolm Atkinson, Université de Glasgow ; Equipe de Ron Morrison, Université de St-Andrews ; Equipe d'Antonio Albano, Université de Pise)

#### Période 1990 à 1998 : Maître de Conférences Université de Paris 1

<sup>1</sup>Formally Integrated Data Environment

#### Thématiques :

Aspects système des bases de données (1986-1990, doctorat), Contraintes d'intégrité dans les bases de données (1991-1999, maître de conférences), typage et sécurité pour données semi-structurées, langages pour données, langages de workflows pour données scientifiques (depuis 2001).  
Encadrement :  
Direction de 9 thèses, de 19 stages de DEA, d'un stage pour normaliens.

## Langages pour bases de données et intégrité

À partir de Novembre 1991, j'ai poursuivi mes activités de recherche dans le domaine de la définition de langages de programmation pour bases de données. Dans ce cadre, j'ai plus particulièrement étudié le problème de la définition et de la vérification de contraintes d'intégrité pour les bases de données orientées objet.

Dans le domaine des bases de données il est crucial de pouvoir spécifier des propriétés (contraintes d'intégrité) que les données stockées dans la base doivent satisfaire (par exemple que toute personne est soit célibataire, soit le conjoint de son conjoint).

La plupart des langages de programmation pour bases de données tels que *Galileo*, *Napier88*, *Fibonacci* ou  $O_2$  ne permettaient pas d'exprimer ce type de propriétés de façon globale et déclarative.

Ma contribution a consisté à étudier la possibilité d'intégrer ce type de fonctionnalités au sein d'un langage de programmation pour bases de données. Trois aspects importants ont été étudiés tant sur le plan formel que pratique : (i) définir un langage de spécification de contraintes d'intégrité, (ii) réduire statiquement (à la compilation) le nombre de contraintes d'intégrité à vérifier lors de l'exécution d'une transaction (d'un programme), et (iii) générer automatiquement des algorithmes de vérification efficaces.

### Thémis

Le premier point nous a conduit à développer un langage (*Thémis*) offrant ce type de fonctionnalités. Une première version de ce langage, proposée dans [27, 26] puis reprise et étendue dans [8], permet d'exprimer de façon globale et déclarative des contraintes d'intégrité qui ne portent que sur des attributs statiques (des données stockées).

### Analyse statique et déduction automatique :

Un premier mécanisme, assez rudimentaire [26, 8] (basé sur des informations relatives au typage des données) permettant de détecter *statiquement* si une transaction ne violera pas une contrainte donnée a été défini pour ce langage.

Ce type de mécanisme de détection statique de la préservation de l'intégrité proposé s'avérant assez grossier, j'ai, en collaboration avec Xavier Schaefer (dans le cadre de son stage de D.E.A., que j'ai dirigé en 1993), utilisé la théorie de l'interprétation abstraite (définie par P et R Cousot en 1976) afin d'obtenir des algorithmes de détection de non violation de l'intégrité plus précis, fiables et implantables. Nous adaptons cette théorie à la problématique des bases de données en considérant des langages de programmation persistants. Les premiers résultats de ces travaux ont été présentés dans [36, 9].

Nous avons, ensuite, approfondi ces résultats. Notre approche consiste à définir un contexte d'analyse de programme fondé sur la notion de transformeurs de prédicats (définis par Dijkstra en 1976) et de ramener le problème de la préservation de l'intégrité à celui de la preuve d'un théorème de logique du premier ordre. Plus précisément, un *transformeur de prédicat avant* est une fonction qui prend comme arguments un programme et une formule satisfaite par les données en entrée du programme (une pré-condition) et donne comme résultat une nouvelle formule (une post-condition) qui est satisfaite par tout état qui résulterait de l'exécution du programme. La contrepartie naturelle d'un transformeur de prédicats avant est la notion de *transformeur arrière* qui génère, à partir d'une post-condition, une pré-condition suffisante sur les données en entrée du programme pour que celui ci préserve la post-condition. Ainsi, si la post-condition générée implique la contrainte d'intégrité ou si la contrainte implique la pré-condition, nous sommes sûrs que le programme préserve la cohérence des données.

Nous avons défini et implanté ces deux transformeurs distincts. Pour le transformeur avant, la méthode de déduction utilisée et également implantée pour prouver l'implication est la méthode des tableaux (car le théorème à prouver est en forme normale disjonctive, la résolution opérant sur des formules en forme normale conjonctives). Pour le transformeur arrière, la méthode de déduction utilisée est la résolution (pour des raisons symétriques).

Ceci permet de déduire, de façon plus précise, qu'un programme donné ne violera pas une contrainte [25, 35, 24, 23]. Les résultats de mesures qui ont été menées au moyen de ce prototype sont tout à fait satisfaisants et démontrent la "faisabilité" d'une telle approche. La thèse de X. Schaefer que j'ai dirigée est le fruit de ce travail.

Une évolution naturelle de notre travail a consisté à l'adapter aux contraintes d'intégrité temporelles. De telles contraintes, contrairement aux précédentes, portent sur une succession d'états de la base de données. Un exemple classique de telles contraintes est donné par : "l'âge d'une personne ne peut décroître". Nous avons défini des mécanismes de détection statique inspirés des techniques précédemment décrites combinés à des techniques de déduction automatique en logique temporelle du premier ordre. Le problème majeur qui s'est posé, dans ce contexte, a été de définir un démonstrateur de théorèmes pour la logique temporelle du premier ordre. Ceci a fait l'objet du stage de DEA de S. Praud et s'est poursuivi, dans le cadre de sa thèse que j'ai dirigée en collaboration avec S. Cerrito (LRI) [7, 34].

#### Génération automatique de tests :

Concernant la génération automatique d'algorithmes de vérification le problème se ramène au problème théorique suivant : étant donné une formule de la logique du premier ordre (la contrainte) et un modèle fini de cette formule (la base de données) si l'on modifie (le programme exécuté) ce modèle fini existe-t-il un moyen constructif efficace pour savoir si la nouvelle interprétation constitue à nouveau un modèle fini de la formule ? Dans le cas général, si l'on ne suppose pas que l'on dispose d'un modèle fini initialement, le problème est PSPACE-complet (réduction de QBF). Une classe de formules a été isolée (formules possédant au plus une alternance de quantificateurs existentiels/universels ou universel/existential) pour lesquelles les algorithmes proposés sont efficaces (polynômiaux) [26, 8]. S. Monties, dans son stage de D.E.A., étend ces algorithmes au cas des quantificateurs existentiels. Initialement écrits dans un style impératif, ces algorithmes ont été réécrits dans le langage de requête standard pour les bases de données orientées-objet : OQL. Ceci a fait l'objet du stage de D.E.A., de L. Tonnelier. Ces algorithmes de vérification sont définis et automatiquement générés lors de la compilation des transactions.

### Période 1998 à ce jour : Professeur Université de Paris 11

#### **Langages de programmation et systèmes pour données semi-structurées**

L'explosion récente d'Internet comme support de calculs globaux conduit à l'échange croissant d'informations. La nature hautement distribuée du réseau induit que les informations sont généralement collectées, traitées puis restructurées à partir de sources multiples et hétérogènes. Afin d'en permettre le traitement ces informations sont publiées sous un même format. XML devient le standard de-facto pour l'échange et la manipulation de documents sur le Web. De nombreux facteurs plaident en faveur de son acceptation : les documents XML sont lisibles par un humain, ils sont par nature auto-descriptifs ; il fait l'objet d'un processus actif de standardisation au sein du W3C et de nombreuses entreprises, parmi les plus représentatives dans le domaine des technologies de l'information, l'ont d'ores et déjà adopté. Dans un tel contexte nous nous intéressons aux points suivants :

1. Définition de langages de programmation pour XML typés et expressifs

## 2. Systèmes de stockage natifs et optimisation de requêtes

### Langages de programmation

Comme le reflète le processus actuel de standardisation, les documents XML sont intrinsèquement typés (notions de validité et de bonne formation relativement à une DTD, XML-Schema, Relax-NG). Or, les langages de transformation associés (XPath/XSLT) sont essentiellement non-typés.

La définition de langages de transformations pour XML typés permettra d'obtenir une plus grande fiabilité du code et surtout d'offrir un contrôle (partiel) statique de correction de la transformation. Ainsi, dans un tel contexte, il sera possible de garantir que le résultat d'une transformation donnée est bien conforme au type (ou DTD, ou Schemas) attendu.

Enfin, dans le contexte de l'interrogation stricto-sensu de documents XML, il est indispensable d'offrir une interface plus déclarative qu'un langage turing-complet. Ce type d'interface est communément appelé langage de requête. Notre contribution des dernières années pour cet axe est brièvement développée.

#### Le Langage CDuce ([www.cduce.org](http://www.cduce.org))

Les langages de transformation de documents XML actuels sont loin d'être complètement satisfaisants. Dans le cas de DOM (Document Object Model) —un API pour XML très répandu— par exemple, il n'y a aucune manière de garantir statiquement que le document produit par un programme est du type attendu. Ceci est aussi vrai pour XSLT par exemple, à l'exception près que dans ce dernier cas des programmes non corrects n'engendrent pas du XML mal formé mais des documents vides. D'une certaine manière on peut considérer ceux-ci comme des langages non typés.

Proposer un langage typé pour manipuler des documents XML est assez naturel : XML n'est qu'une syntaxe, et différents systèmes de types sont proposés : DTD, XML-Schema, ... ; un document ne prend de sens que si on le voit avec un certain type XML.

Dans le cadre du stage de DEA de A. Frisch que j'ai co-encadré avec G. Castagna nous avons étudié l'introduction du sous-typage sémantique ceci nous a permis d'établir le socle théorique sur lequel bâtir les développements futurs. Les résultats de ce travail sont présentés dans [20, 6] et donnent lieu au développement du langage CDuce [19, 18] qui est disponible sur le site [www.cduce.org](http://www.cduce.org) et qui compte à ce jour une centaine d'utilisateurs et est intégré aux distributions Linux (Ubuntu/Debian, Mandriva, Fedora).

Nous poursuivons ces recherches dans deux directions parallèles : (i) définition de langages de requêtes et d'interfaces graphiques pour les données au format XML et (ii) stockage et optimisation pour l'accès aux documents XML.

#### ⒸQL un langage de requête pour XML:

Le système de types de XDuce sert de base à celui de XQuery, une algèbre de requêtes pour XML proposée par le W3C. L'adoption d'un tel système montre l'intérêt que porte le monde des bases de données à celui des langages typés adaptés à XML. Cet intérêt ne se limite pas uniquement au monde académique. En effet, les types peuvent apporter plusieurs caractéristiques au monde des bases de données (comme ils l'ont toujours fait de part le passé).

Dans le cadre de la manipulation de documents XML, on peut considérer le mécanisme de filtrage des langages fonctionnels de la famille d'OCaml ("pattern-matching") comme une primitive pour exprimer et exécuter une *requête*. Dans le contexte du stage de Cédric Michon [33], puis de sa thèse en cours [16], nous avons suivi cette approche en intégrant dans CDuce un langage de requête, ⒸQL, exploitant pleinement la puissance expressive des

motifs (“patterns”) ainsi que leur compilation efficace au moyen d’automates d’arbres non-uniformes. Il s’agit d’un langage déclaratif qui laisse toute latitude au compilateur pour exploiter au mieux des réécritures algébriques (en particulier de motifs) afin de générer des plans d’exécutions, en mémoire centrale efficaces.

Les perspectives de ce travail sont diverses: (i) étendre l’opération de filtrage afin de permettre également des parcours en profondeur de l’arbre (à la XPath), et (ii) offrir un mécanisme de persistance à CQL ainsi que des mécanismes d’optimisation physique de requête associés.

#### Pattern by Example (PBE) : une interface graphique pour XML:

Il est commun que les utilisateurs d’un système de bases de données ne soient pas experts. Naturellement, le système offre une interface graphique pour permettre à de tels utilisateurs d’interroger et d’extraire de l’information. C’est pourquoi un des prolongements du travail effectué sur CQL a été de définir un langage de requêtes graphique [10].

Cette interface s’inspire du langage graphique QBE (Query by Example) défini pour les systèmes relationnels et l’adapte au cas de XML. Dans un tel contexte, le fait de définir un tel langage est crucial car non seulement la nature des données est beaucoup plus complexes mais en outre, les langages de manipulation sont basés sur le mécanisme de “pattern-matching” des langages fonctionnels qui sont difficiles à appréhender pour un utilisateur lambda. Ainsi, l’utilisateur interagit avec le système au travers de cette interface et la requête correspondante est automatiquement générée dans une forme optimisée. La génération est guidée par les types et la preuve formelle de la correction de la traduction est établie.

### **Données XML : aspects systèmes**

#### Stockage natif et optimisation de requêtes

Les résultats existants en matière d’optimisation de requêtes sur des données (XML) persistantes peuvent être classés en fonction du modèle sous-jacent de stockage et de traitement de requêtes. Des travaux se sont appuyés sur des systèmes de gestion de bases de données relationnelles pour stocker des documents XML, et ont traité des requêtes XML en les traduisant en SQL. En conséquence, tous les aspects reliés au placement des données, au stockage, et à l’optimisation de requêtes ont été délégués au système relationnel. Cette approche a ses limites, puisque des langages de requêtes XML plus récents (tels que XQuery et CDuce sont bien plus complexes, et très différents de SQL.

Plus généralement, en collaboration avec Ioana Manolescu (INRIA - Futurs - Gemo), avec qui j’ai co-encadré la thèse d’A. Arion, nous proposons un optimiseur générique [14, 15, 12, 11] pouvant être paramétré en fonction des besoins et caractéristiques des applications visées.

Parallèlement, dans le cadre du travail de thèse de K. NGuyen, nous avons défini des techniques d’optimisation de requêtes (exprimées en XQuery) qui permettent grâce à une analyse statique de la requête et à la connaissance du type du document à interroger, d’élaguer ce dernier pour ne charger en mémoire centrale que la portion nécessaire à l’évaluation de la requête. Ce travail a été publié à VLDB [13]

### **Langages de workflow interactifs pour données scientifiques massives et distribuées**

Les applications de gestion de données scientifiques sont, de nos jours, amenées à gérer d’énormes volumes de données dont la taille croît sans cesse. Ces données peuvent être de nature très diverse : numériques (résultats d’instruments de mesure), textuelles (corpus étudiés par des chercheurs en sciences humaines) structurées (en astronomie ou en physique) ou totalement irrégulières (dossiers médicaux).

Quelle que soit leur nature, ces données sont caractérisées par leur taille importante. Ceci résulte, d'une part, de l'utilisation de plus en plus répandue d'ordinateurs pour recueillir ces informations en lieu et place des cahiers utilisés auparavant par les chercheurs et, d'autre part, de l'utilisation d'outils puissants et automatisés pour intégrer de telles informations par exemple des télescopes, des navigateurs ciblés sur tel ou tel domaine ...

Ces données représentent une mine d'or pour les scientifiques qui peuvent fonder leurs recherches sur une telle base d'information. La disponibilité de ces données est déterminante quant à leur exploitation par une communauté scientifique. Or, le plus souvent les scientifiques utilisent, pour stocker ou manipuler de telles sources : des formats propriétaires, des systèmes ad-hoc, des processus "maison", ce, clairement au détriment de la disponibilité et donc du partage à grande échelle.

Par ailleurs, au sein de la communauté informatique, de nombreux efforts de recherche portent, indépendamment, sur la visualisation d'informations, la gestion de processus et la gestion de données massives. Ces trois courants nous semblent indispensables à l'obtention de bonnes plateformes de gestion de données scientifiques permettant d'aider les chercheurs dans leur travail quotidien. En effet, les propriétés suivantes caractérisent la nature des applications de gestion de données scientifiques : la complexité calculatoire et le volume massif des données à traiter, l'interactivité des processus mis en jeu et la visualisation indispensable des résultats. Par conséquent, nous identifions les trois thématiques suivantes :

1. Visualisation d'information et workflows interactifs,
2. Bases de données
3. Gestion de processus et systèmes de workflow

Ces trois thématiques sont complémentaires. Si les systèmes de visualisation d'informations permettent de présenter les données de façon claire, les algorithmes et outils qu'ils utilisent sont trop complexes pour un utilisateur naïf. Les systèmes de workflow interactifs, quant à eux, ne sont pas en mesure de faire face à la taille colossale des données scientifiques. D'autre part, les systèmes de gestion de bases de données manquent de flexibilité (les modifications de schéma sont difficiles), de puissance calculatoire (les langages de requêtes ne sont pas complets au sens de Turing) et enfin ils offrent, le plus souvent, des interfaces très rudimentaires. En revanche, par la nature déclarative de leurs langages de requêtes sous-jacents et par l'état de maturité atteint par leurs techniques d'optimisation, de gestion des accès concurrents et de stockage à grande échelle, ils possèdent de nombreux atouts. C'est ainsi qu'à la confluence des bases de données et du génie logiciel, nous trouvons le domaine de la gestion de (business) processus et de "workflows". Or, ce cadre est trop ciblé sur des applications de gestion de processus métiers et les systèmes développés par Oracle, SAP ou IBM ne sont pas adaptés à la gestion de données scientifiques. Ils sont complexes à utiliser et requièrent un administrateur spécifique, leur coût est également prohibitif. En outre, ils sont adaptés au traitement de données représentées sous forme tabulaire ce qui n'est pas le cas des données scientifiques qui ont en général une structure plus riche. Enfin ils ne permettent pas d'opérer des calculs puissants indispensables en revanche dans le domaine d'applications scientifiques.

Récemment, des plateformes de gestion de données scientifiques émergent outre-atlantique (Kepler/Ptolemy). Cependant elles ne prennent pas en compte la nature interactive des processus liés au traitement des données scientifique ni n'offrent des outils puissants de visualisation des résultats. Or les deux aspects sont indispensables à l'obtention de systèmes adaptés aux besoins des chercheurs.

Dans le cadre de la thèse de Wael Khemiri que nous co-encadrons, depuis septembre 2008, avec Ioana Manolescu, nous proposons de *définir des modèles, des algorithmes ainsi que des architectures* en vue de l'obtention d'un système en mesure d'apporter aux scientifiques les moyens de tirer le meilleur parti des données qu'ils recueillent en leur permettant de les organiser à leur guise. Ce travail de recherche s'inscrit dans le cadre du projet Digitéo EDIFLOW pour lequel collaborent trois équipes : Gémo (I. Manolescu, INRIA Saclay) AVIZ (J-D Fekete,

INRIA - Saclay) et l'équipe Bases de Données du LRI. Il met en jeu trois différents domaines desquels nous espérons prendre le meilleur.

## Encadrement.

Le pourcentage de participation à l'encadrement est indiqué entre parenthèses carrées.

### Thèses

1. Wael Khemiri, "Workflow Interactifs pour données scientifiques", depuis Septembre 2008 [50%]
2. Kim N'Guyen, "Algèbres de filtrage pour l'interrogation de documents XML", 2004-2008 (thèse soutenue en Mai 2008) [50%]
3. Andrei Arion, "Optimisation à base d'index et mises à jour dans un système de gestion de données XML", 2004-2007 [50%].
4. Cédric Miachon. *Requêtes dans CDuce : implantation, optimisation et mesures de performances*, 2003-2006.
5. Marwan Burelle. *Sécurité de transformations XML* (depuis septembre 2002) Co-encadrement avec G. Castagna, LIENS, ENS [50%]. (Abandon pour raisons familiales).
6. S. Praud 1997-2000. Co-encadrement avec S. Cerrito (LRI)  
Sujet : "Contraintes temporelles, analyse statique et déduction automatique". Thèse soutenue en Février 2002, Université de Paris XI. [50%]
7. F. Watez 1997-2000 co-encadrement avec S. Cluet (Projet Verso-INRIA).  
Sujet : "Optimisation de requêtes OQL".  
Thèse effectuée sous convention CIFRE (Société O<sub>2</sub>Technology), Thèse soutenue en Novembre 2000. [50%]
8. X. Schaefer, 1994-1997  
Sujet : "Bases de données orientées objet, Contraintes d'intégrité et Analyse statique".  
Thèse soutenue le 15 Octobre 1997, Université de Paris 1 - Panthéon - Sorbonne.[100%]
9. P. Y. Policella de 1992 à 1994 co-encadrement en collaboration avec A. Doucet. Thèse soutenue le 11 Juin 1996, Université d'Orsay.  
Sujet : "Cohérence dans les bases de données orientées objet". [50%]

### DEA - Master Recherche

1. Elise Henaux, "Validation sémantique de données au format XML pour la norme CGNS", Année 2004-2005
2. Kim N'Guyen, "Algèbres de filtrage pour l'interrogation de documents XML", Année 2003-2004 [50%]
3. Andrei Arion, "Optimisation à base d'index et mises à jour dans un système de gestion de données XML", Année 2003-2004 [50%].
4. Cédric Miachon. "Requêtes dans CDuce : implantation, optimisation et mesures de performances", Année 2002-2003.

19 dea - master recherche,  
9 thèses ; Participation à 25  
jurys de thèse ou d'HDR.

5. Marwan Burelle. “Sécurité de transformations de documents XML” DEA I<sub>3</sub>, Année 2001-2002 [50%]
6. Alain Frisch. “Types récursifs, combinaisons booléennes et fonctions surchargées: application au typage de XML” DEA Programmation. Année 2000-2001. [50%]
7. M. Pasquier (98-99)  
Sujet : “Vérification statique d’assertions en java”.
8. A. Diallo (98-99)  
Sujet : “étude de quelques calculs pour la mobilité”.
9. H. Benajam (97-98)  
Sujet : “Intégrité et test de logiciel”.
10. S. Praud (96-97)  
Sujet : “Contraintes d’intégrités temporelles et analyse statique de programmes”.
11. R. Zhang (96-97)  
Sujet : “Analyse statique du langage DEL : une application à la gestion de l’intégrité pour le système Validity”.
12. N. Hugon (95-96) en collaboration avec E. Pelz.  
Sujet : “Vers une sémantique pour les règles actives via le Petri-box calculus”.
13. L. Tonnelier (94-95)  
Sujet : “Contraintes d’intégrité et OQL: vers une intégration de contraintes d’intégrité dans le système O<sub>2</sub>”.
14. M. Sarkis (93-94)  
Sujet : “Contraintes d’intégrité dans les bases de données orientées objet : approche procédurale versus déclarative.”
15. R. Souah (93-94)  
Sujet : “Migration d’objets en présence de contraintes d’intégrité”.
16. S. Montiès (92-93)  
Sujet : “Algorithmes de vérification de contraintes d’intégrité fondés sur des méthodes de simplification : application aux quantificateurs existentiels”.
17. X. Schaefer (92-93)  
Sujet : “Interprétation abstraite de langages persistants : une application à la satisfaction de contraintes d’intégrité”.
18. S. Attali (91-92)  
Sujet : “Contrôle d’intégrité et méthodes de vérification”
19. P. Y. Policella (91-92)  
Sujet : “Etude et définition de contraintes d’intégrité dans un environnement orienté-objet”.

### **Jurys et rapports de thèses**

1. Membre du jury de l’Habilitation à diriger des recherches de Ioana Manolescu, Université de Paris 11, octobre 2009.
2. Membre du jury de la thèse de Bogdan Marinoiu, Université de Paris 11, Juin 2009.

- 
3. Membre du jury de la thèse de V. Sans, Université de Cergy, Décembre 2008.
  4. Membre du jury de la thèse de M. Sozeau, Université de Paris 11, Décembre 2008.
  5. Membre du jury de la thèse de N. Preda, Université de Paris 11, Mai 2008.
  6. Membre du jury de l'Habilitation à diriger des recherches de Mirian Halfeld, Université de Blois, décembre 2007.
  7. Membre du jury de la thèse de N. Travers, soutenue le 9 décembre 2006 à l'Université de Versailles Saint Quentin.
  8. Membre du jury de la thèse de L. Pilard, soutenue le 15 décembre 2005 à l'Université de Paris 11.
  9. **Rapporteur** de la thèse de D. Debarbieux, soutenue le 9 décembre 2005 à l'Université de Lille.
  10. **Rapporteur** de la thèse de S. Medjdoub, soutenue le 8 décembre 2005 à l'Université de Versailles Saint Quentin.
  11. Membre du jury de la thèse de N. Gesbert soutenue le 2 décembre 2005 à l'Université de Marne la Vallée.
  12. **Rapporteur** de la thèse de D. Duarte, soutenue en juillet 2005 à l'Université de Francois Rabelais Blois.
  13. Membre du Jury de la thèse d'O. Benjelloun, soutenue en septembre 3004 à l'Université de Paris 11.
  14. Membre du Jury de la thèse de P. Letouzey, soutenue le 10 Juillet 2004 à l'Université de Paris 11.
  15. Membre du Jury de la thèse de J. Forest, soutenue le 29 Septembre 2003 à l'Université de Paris 11.
  16. Membre du Jury de la thèse de P-A. Veltri, soutenue le 1 Octobre 2002 à l'Université de Paris 11.
  17. Membre du Jury de la thèse de J. Barbay, soutenue le 24 Septembre 2002 à l'Université de Paris 11.
  18. **Rapporteur** de la thèse de L. Nemirovski, soutenue à l'Université de Paris 1 Janvier 2002.
  19. Membre du jury de la thèse de F. Watez, Soutenu à l'Université de Paris 1 le 31 Octobre 2000.
  20. Membre du jury de la thèse de D. Dvekausas, Soutenu à l'Université de Paris XI le 15 Juin 1999.
  21. **Rapporteur** de la thèse de A. Ait-Bouziad, Soutenu à l'Université de Paris VII, le 1 Juin 1999.
  22. **Rapporteur** de la thèse de J. Darmont, Soutenu à l'Université de Clermond-Ferrand, le 18 Janvier 1999.
  23. **Rapporteur** de la thèse de F. Bentayeb, Soutenu à Université d'Orléans en Juin 1998,

24. Membre du jury de la thèse de F. Semmak soutenue à l'Université de Paris 1 le 24 Février 1998.
25. Membre du jury de la thèse de P.Y. Policella soutenue à l'Université d'Orsay le 11 Juin 1996.

#### Direction de stages pour élèves normaliens

- E.-O. Lebigot, G. Lecerf. "Génération de Bases Abstraites à Partir de Définitions de Type et d'une Contrainte Logique". Année 1994-1995. [50%]

#### Prime d'encadrement doctoral

**Titulaire de la prime d'encadrement doctoral** d'Octobre 1993 à Novembre 2001 puis depuis Octobre 2006.

## Évaluation et administration de la recherche.

#### Jurys de concours

1. Septembre 2009, Membre de la commission pour la prime d'excellence scientifique 27ème section.
2. Mai 2009, Membre du comité de sélection poste de Maître de Conférences, IUT, Paris 11.
3. Mai 2009, Membre du comité de sélection poste de Professeur, INPG
4. Mai 2008, Membre du jury du concours CR2 pour l'Inria Saclay.
5. Mai 2008, Membre du jury du concours CR2 pour l'Inria de Rennes.
6. Mai 2007 Membre des jurys du concours CR2 et CR1 pour l'Inria de Rennes.
7. Mai 2005, Membre du jury du concours CR2 pour l'Inria de Rennes.
8. Mai 2004, Membre du jury du concours CR2 pour l'Inria de Rennes.
9. Mai 2004, Membre du jury du concours CR2 pour l'Inria de Futurs.
10. Mai 2003, Membre du jury du concours CR2 pour l'Inria de Rennes.

#### Comités de rédaction

1. Membre du comité de rédaction de la revue  $I_3$ , depuis 1999.
2. Membre du comité de rédaction de la revue TSI, de Janvier 1996 à Janvier 2000.
3. Membre du comité de lecture du numéro spécial de la revue L'Objet, "Méthodes formelles pour les systèmes à objets", 1999.

Membre de jurys de concours INRIA, Présidente du Comité de programme de BDA'2005, Membre de plus de 18 comités de programme. Coéditeur d'un numéro spécial de Techniques et Sciences Informatiques; coéditeur des actes du 4ème International Workshop on Persistent Object Systems, membre du comité de programme du workshop international Xime-P, membre de comités de sélection de conférences, workshops

Organisatrice de journées et ateliers dont 2 internationaux. Lecteur pour nombreuses revues et conférences internationales.

---

**Comités de programme**

1. Membre du comité de programme de la conférence internationale VLDB, 2009.
2. Membre du comité de programme de l'ACM International workshop on Programming Languages for XML (PLAN-X 2008).
3. Membre du comité de programme de la conférence Bases de données avancées (BDA 2007).
4. Membre du comité de programme de l'ACM - International workshop on Xquery Implementation and Experience XIME-P 2007.
5. Membre du comité de programme de l'International XML Symposium (XSYM 2007)
6. Membre du comité de programme de l'International XML Symposium (XSYM 2006)
7. Membre du comité de programme de l'International XML Symposium (XSYM 2005)
8. Membre du comité de programme de l' International Workshop on Programming Language technologies for XML (PLAN-X 2005) in conjunction with ACM POPL'2005.
9. Membre du comité de programme de l' International Symposium on Methodologies for Intelligent Systems 2005
10. **Présidente du comité de programme** de la conférence Bases de Données Avancées (BDA) 2005.
11. Membre du comité de programme du workshop on XQuery, Xime-P joint with ACM SIGMOD-PODS, Paris Juin 2004.
12. Membre du comité de programme de la conférence Bases de Données Avancées (BDA) 2003.
13. Membre du comité de programme de la conférence Bases de Données Avancées (BDA) 2002.
14. Membre du comité de programme de la conférence Bases de Données Avancées (BDA) 1999.
15. Membre du comité de programme de la conférence Bases de Données Avancées (BDA) 1996.
16. **Co-chair** du comité de programme de International Workshop on Persistent Object Systems (POS6), 1994.
17. Membre du comité de programme de la conférence Bases de Données Avancées (BDA) 1993.
18. Membre du comité de programme de la conférence Inforsid, 1992.

**Comités d'organisation**

1. **Co-responsable** de l'organisation de "Persistent Object Systems 6" (POS6), 5-9 Septembre 1994, Tarascon, France.
2. Participation à l'organisation du "2nd Workshop on Foundation of Object-Oriented Languages", Paris juillet 1994.
3. Organisation du meeting des partenaires francophones du réseau d'excellence RENOIR, Décembre 1997, Paris.

## Autres

- Lecteur pour les revues Journal of Functional Programming, Information Systems, Theoretical Computer Science, Theory and Practice of Object-Oriented Systems et pour les conférences DBPL, PODS, VLDB, EDBT, ICDT, ICALP, ESOP, ECOOP.

## Conférences et séminaires invités

La personne qui m'a invité est indiquée entre parenthèses carrées.

“Systèmes de Bases de données orientées-objet”, Tutorial invité. Conférence Inforsid 1991.

“Static verification of integrity constraints in object-oriented database programming languages via predicate transformers”, Séminaire invité, Sémantique et Interprétation Abstraite [P. Cousot], École Normale Supérieure, Avril 1997.

“Forward and backward analysis of database programming languages”, Séminaire invité, [J. Vitek] Université de Genève, Suisse, Septembre 1997.

“Forward and backward analysis of database programming languages”, Séminaire invité, [A. Carbone] Université de Paris 12, Mai 1998.

## GDR Bases de Données

1. De 1995 à 1998 responsable du serveur Web du GDR BD3.
2. De 1995 à 1998 modérateur de la “mailing list” du GDR BD3.

## Animation de la recherche

1. De Septembre 2002 à la fin des RTP, membre du comité de pilotage du RTP Bases de données et d'informations hétérogènes et distribuées (RTP 9).
2. Janvier 2003, rapporteur extérieur pour l'évaluation d'une proposition RNTL.
3. Octobre 2006 rapporteur d'une proposition pour le Netherlands Organisation for Scientific Research (NWO).

## Contrats et collaborations

**1989-1992:** membre du projet ESPRIT BRA n° 3070 FIDE.

**1997-1999** membre du réseau d'excellence RENOIR (Requirement Engineering Network On International cooperating Research groups).

**1999-2000** Participation au groupe “Spécification, vérification, sémantique”, pôle *Objets et composants logiciels*, GDR ALP (1999).

**2003-2004** Porteur d'un projet d'action spécifique “Langages, types, sécurité et intégrité pour données semi-structurées”, dans le cadre du RTP 9 acceptée en mai 2003 (<http://www.lri.fr/~benzaken/ltsi.html>).

**2003-2006** Responsable pour l'équipe Bases de donnée d'un proposition soumise à l'appel d'offre RNTL : “GraphDuce”, labellisée le 13 mars 2003.

**2003-2006** Responsable pour le partenaire LRI, du projet CASC, ACI-sécurité du CNRS-MENRT-INRIA.

**2004-2008** Responsable pour le partenaire LRI, du projet TRALALA, ACI-masse de données du CNRS-MENRT-INRIA.

**2008-2011** membre du projet ANR - Codex

**2008-2011** porteur pour l'équipe Bases de données du LRI du projet Digitéo Ediflow (Efficient Interactive Workflow for Scientific Data Management).

## Valorisation, Transfert.

- Réalisation du module de regroupement d'objets sur disque pour le système de gestion de bases de données commercial O<sub>2</sub>.
- Le langage CDuce ainsi que son implantation font l'objet d'un dépôt de dossier (en cours) de propriété intellectuelle auprès du service de valorisation du CNRS.
- CDuce est **intégré aux distributions Linux** suivantes : Ubuntu/Debian, Fedora et Mandriva.

## Publications.

### Livres

- [1] V. Benzaken and A. Doucet. *Bases de données orientées objet : origines et principes*. Armand-Colin, 1993.

### Chapitres de livres

- [2] V. Benzaken, G. Castagna, H. Hosoya, B.C. Pierce, and S. Vansummeren. *The Encyclopedia of Database Systems*, chapter "XML Typechecking". Springer, 2009. Article invité.
- [3] V. Benzaken, C. Delobel, and G. Harrus. Clustering strategies in o<sub>2</sub>: an overview. In F. Bancilhon, C. Delobel, and P. Kanellakis, editors, *Building an Object-Oriented Database System: the Story of O<sub>2</sub>*, pages 385–410. Morgan Kaufman, 1992.

### Édition d'ouvrages collectifs

- [4] M. Adiba and V. Benzaken, editors. *Numéro spécial Bases de données*, volume 17 of *TSI*. Hermès, 1998.
- [5] M. Atkinson, V. Benzaken, and D. Maier, editors. *Persistent Object Systems 6*. Workshops in Computing. Springer-Verlag, Tarascon, France, 5-9 September 1994.

### Articles dans des revues internationales avec comité de lecture

- [6] Alain Frisch, Giuseppe Castagna, and Véronique Benzaken. Semantic subtyping: Dealing set-theoretically with function, union, intersection, and negation types. *Journal of the ACM*, 55(4):1–64, 2008.
- [7] V. Benzaken, S. Cerrito, and S. Praud. Static verification of dynamic constraints: a semantics-based approach. *Network and Information Systems Journal (NIS)*, 2(5-6):549–571, 1999.
- [8] V. Benzaken and A. Doucet. Thémis: A Database Programming Language Handling Integrity Constraints. *Very Large Databases (VLDB) Journal*, 4(3):493–518, 1995.

Un livre (Armand-colin); un chapitre de livre; coéditeur d'actes dans la série "workshop in computing" Springer-Verlag, coéditeur d'un numéro spécial de *TSI*; 3 Revues internationales de forte audience, 1 revue nationale, 21 conférences internationales

---

**Articles dans des revues nationales  
avec comité de lecture**

- [9] V. Benzaken, A. Doucet, and X. Schaefer. Integrity Constraint Checking Optimization based on Abstract Databases Generation and Program Analysis. *Journal de l'Ingénierie des Systèmes d'Information*, 1(3):9–29, 1995.

**Communications dans des conférences internationales avec comité de lecture**

- [10] V. Benzaken, G. Castagna, D. Colazzo, and C. Miachon. Pattern by example: Type-driven visual programming of XML queries". In *10th International ACM SIGPLAN Symposium on Principles and Practice of Declarative Programming*, Valencia, Spain, July 2008. ACM.
- [11] A. Arion, V. Benzaken, I. Manolescu, and Y. Papakonstantinu. Structured materialized views for XML queries. In *33rd International Conference on Very Large Databases (VLDB 2007)*, September 2007.
- [12] A. Arion, V. Benzaken, I. Manolescu, and Y. Papakonstantinu. Algebra-based identification of tree patterns in xquery. In Henrik Legind Larsen, Gabriella Pasi, Daniel Ortiz-Arroyo, Troels Andreasen, and Henning Christiansen, editors, *Flexible Query Answering Systems: 7th International Conference, FQAS 2006*, volume 4027 of *LNCS*, pages 13 – 25, Milano, June 2006. Springer Verlag.
- [13] V. Benzaken, G. Castagna, D. Colazzo, and K. NGuyen. Type-based XML projection. In *32nd International Conference on Very Large Databases (VLDB 2006)*, September 2006.
- [14] Andréi Arion, Véronique Benzaken, and Ioana Manolescu. XML access modules: Towards physical data independence in XML databases. In *Proceedings of the Second International Workshop on XQuery Implementation, Experience and Perspective colocated with ACM Sigmod Pods 2005*, 2005.
- [15] Andréi Arion, Véronique Benzaken, Ioana Manolescu, and Ravi Vijay. Uload: Choosing right storage for your XML application. In *Demonstration at the 31st International Conference on Very Large Databases*, 2005.
- [16] Véronique Benzaken, Giuseppe Castagna, and Cédric Miachon. A full pattern-based paradigm for XML query processing. In *PADL, International Symposium on Practical Aspects of Declarative Languages*, pages 235–252, 2005.
- [17] V. Benzaken, M. Burelle, and G. Castagna. Information flow security for XML transformations. In V. Saraswat, editor, *Eight Asian Computing Science Conference (ASIAN'03)*, Lecture Notes in Computer Science, Mumbai, India, 10-13 December 2003. Springer-Verlag.
- [18] V. Benzaken, G. Castagna, and A. Frisch. CDuce: an XML-centric general-purpose language. In *ACM International Conference on Functional Programming (ICFP'03)*, 2003.
- [19] V. Benzaken, G. Castagna, and A. Frisch. CDuce: a white paper. In B. C. Pierce and P. Wadler, editors, *International workshop on Programming Languages for XML (PLANX'02)*, Pittsburgh, USA, 2002.
- [20] A. Frisch, G. Castagna, and V. Benzaken. Semantic subtyping. In *Seventeenth Annual IEEE Symposium on Logic in Computer Science (LICS'2002)*, pages 137–146, Copenhagen, Denmark, 2002.

- 
- [21] F. Watez, S. Cluet, V. Benzaken, and C. Fiegel. Benchmarking queries over trees: Learning the hard truth the hard way (short paper). In *ACM SIGMOD International Conference*, Dallas, May 2000. ACM.
- [22] V. Benzaken, N. Hugon, H. Klaudel, E. Pelz, and R. Rieman. A mnet calculus-based semantics for triggers. In Springer Verlag, editor, *International Conference on Theory and Applications of Petri Nets*, LNCS, Lisbon, Portugal, June 1998.
- [23] V. Benzaken and X. Schaefer. Static management of integrity in object-oriented databases: Design and implementation. In Hans-Jörg Schek, editor, *6th International Conference on Extending Database Technology (EDBT'98)*, Valencia, Spain, March 23-27 1998. Springer Verlag.
- [24] V. Benzaken and X. Schaefer. Static integrity constraint management in object-oriented database programming languages via predicate transformers. In Aksit and Matsuoka, editors, *European Conference on Object-Oriented Programming (ECOOP'97)*, number 1024 in *Lecture Notes in Computer Science*, pages 60–84. Springer-Verlag, 1997.
- [25] V. Benzaken and X. Schaefer. Ensuring efficiently the integrity of persistent object systems via abstract interpretation. In S. Nettles and R. Connor, editors, *Proceedings of Persistent Object Systems 7*, pages 72–87, Cape May, U.S.A, June 1996. Morgan Kaufman.
- [26] V. Benzaken and A. Doucet. Thémis: a database programming language with integrity constraints. In C. Beerli, A. Ohori, and D. Shasha, editors, *Proceedings of Database Programming Languages 4*, Workshop in Computing, pages 243–262, New York City, USA, September 1993. Springer-Verlag.
- [27] V. Benzaken, C. Lécluse, and P. Richard. Enforcing Integrity Constraints in Database Programming Languages. In R. Morrison A. Albano, editor, *Proceedings of Persistent Object Systems 5*, Workshop in Computing, pages 282–299, San Miniato, Italy, September 1-5 1992. Springer-Verlag.
- [28] V. Benzaken. An Evaluation Model for Clustering Strategies in the O<sub>2</sub> Object-Oriented Database System. In S. Abiteboul and P. Kanellakis, editors, *International Conference on Database Theory (ICDT)*, number 470 in *Lecture Note in Computer Science*, pages 126–140, Paris, France, December 1990. Springer Verlag.
- [29] V. Benzaken and C. Delobel. Enhancing Performance in a Persistent Object Store: Clustering Strategies in O<sub>2</sub>. In *Proceedings of Persistent Object Systems 4*. Morgan Kaufmann, September 1990.
- [30] F. Bancilhon, G. Barbedette, V. Benzaken, C. Delobel, S. Gamerman, C. Lécluse, P. Pfeffer, P. Richard, and F. Velez. The design and implementation of o<sub>2</sub>, an object-oriented database system. In *Advances in Object-Oriented Database Systems*, *Lecture Notes in Computer Science*. Springer-Verlag, September 1988.

**Communications dans des conférences nationales  
avec comité de lecture**

- [31] Ioana Manolescu, Wael Khemiri, Véronique Benzaken, and Jean-Daniel Fekete. Reactive workflows for visual analytics. In *Bases de données avancées – BDA*, Namur, 2009.
- [32] I. Manolescu, V. Benzaken, A. Arion, and Y. Papakonstantinou. Structured materialized views for XML queries. In D. Laurent, editor, *BDA 2006*, Lille, October 2006.

- [33] V. Benzaken, G. Castagna, and C. Miachon. CQL a pattern-based query language for XML. In *In Proceedings 20th Bases de Données Avancées (BDA'04)*, 2004.
- [34] V. Benzaken, S. Cerrito, and S. Praud. Static verification of dynamic constraints: a semantics-based approach. In *Actes de la 15ième Conférence sur les Bases de Données Avancées*, Bordeaux, France, 1999.
- [35] V. Benzaken and X. Schaefer. Ensuring efficiently the integrity of persistent object systems via abstract interpretation. In *Actes de la 12ième Conférence sur les Bases de Données Avancées*, Cassis, August 1996. La version étendue se trouve dans [25].
- [36] V. Benzaken, A. Doucet, and X. Schaefer. Integrity constraint checking optimization based on abstract interpretation. In *Actes de la 10ième Conférence sur les Bases de Données Avancées*, Clermont-Ferrand, France, August 1994.
- [37] V. Benzaken, A. Doucet, and P.Y. Policella. Définition et gestion de contraintes d'intégrité dans le langage Thémis. In *Actes de la 9ième Conférence sur les Bases de Données Avancées*, Toulouse, France, October 1993.
- [38] V. Benzaken and A. Doucet. Enforcement Tests Generation for Integrity Constraints Checking Based on Simplification Methods in Object-Oriented Database Systems. In *Actes de la 8ième Conférence sur les Bases de Données Avancées*, Trégastel, France, September 15-18 1992.
- [39] V. Benzaken. Un modèle d'évaluation de stratégies de regroupement dans un système de bases de données orienté-objet. In *Actes de la 6ième Conférence sur les Bases de Données Avancées*, Montpellier, France, September 1990.
- [40] V. Benzaken and C. Delobel. Regroupement d'objets sur disque dans un système de bases de données orienté-objet. In *Actes de la 5ième Conférence sur les Bases de Données Avancées*, Genève, Suisse, September 1989.
- [41] V. Benzaken, C. Delobel, and J.B. N'dala. Gestionnaires de mémoires et d'objets. In *Actes de la 4ième Conférence sur les Bases de Données Avancées*, Bénodet, France, June 1988.

### Thèses

- [42] V. Benzaken. Langages de programmation pour bases de données et contraintes d'intégrité. PhD thesis, Université de Paris-Sud, 1996.
- [43] V. Benzaken. Regroupement d'objets sur disque dans un système de bases de données orienté-objet. PhD thesis, Université de Paris-Sud, 1990.

### Notes de cours

- [44] V. Benzaken and G. Castagna. Langages de programmation pour bases de données. Technical report, DEA théorie et ingénierie des bases de données, Universités Paris I et Paris XI, 1995.

## **Langues Vivantes.**

Langues couramment parlées : Anglais, Français (langue maternelle), Italien.  
Langues parlées : Espagnol (niveau lycée).

*Fait à Paris, le 16 juin 2009*

*Signature*