Start Page

MARIE SKODOWSKA-CURIE ACTIONS

Individual Fellowships (IF)
Call: H2020-MSCA-IF-2015

PART B

"SATIR"

"Static Analysis via Type InfeRence"

This proposal is to be evaluated as:

[Global Fellowship (GF)]

Contents

List of Participants

1 Summary

Software is becoming increasingly complex and critical (autonomous vehicles, Internet, financial systems, etc.). Bugs, security vulnerabilities, or inefficient resource management are costly. Type systems are a scalable way to analyze software but lake of diversity regarding the considered properties. "Graded type systems" will improve the state-of-the-art by including more information in the types and by taking inspiration in abstract interpretation's community. Abstract interpretation is an other software-analyzing tool which analyze is extremely diverse, but which scalability is hindered by a lake of compositionality.

Functional programming languages often use complex type systems describing the behavior of the programs. These types are statically inferred before the compilation, acting as a first test for the correctness of the program. The advantage of type systems is to be inherently compositional and to provide certificate. This means that the libraries comes with their typed that has already been computed and that is easily checkable. However, the verification power is low level and basically only verify that the composition of programs never fails due to incompatibility of format.

More refined static analysis are generally performed by implementations of some abstract interpretation (AI) over the whole program. Those techniques is very efficient but fail to be composable. This means that any new analysis has to be performed through the whole program, including libraries. Such a global analysis has two flaws: is often too resource consuming to be used on a regular basis, and the behaviors of common libraries are approximated while it could be interesting to manually optimize their interpretation once and for all.

In parallel, several functional languages such as Coq, Agda, F* or Idris offer to join to program some proofs of correctness encoded in a dependent types. These dependant type systems are extremely expresive: not only can they cach up to any abstract interpretation analises but they can describe many more refined behaviours. Moreover, these systems are still compositional and with certificates. However, the automatic inference is hardly never automatisable and the programer has to performe the proof himself (helped by the computer).

Notice, also, that performing statical analysis under richelly typed languages should be advantageous: advanced type systems offer several ways to encode semantical information inside the type (for readability or for safety). However, actual tools are basically forgetting the type of a program before running the analysis, loosing all this semantical information given by the programmer.

In this project, we propose to investigate a way to recover as much information as possible from the type inference which is one of the most complex kind of compositional and statical analysis. The ideal outcome would be a static analyzer for functional languages that can compete with existing ones while being completely compositional.

For this purpose, we are focusing on emerging type systems that attach to each type some parameters. Those parameters represent either qualitative statement or quantitative information over resources. Such type systems will be called graded types systems.

A good comparison would be Hoare triplets with preconditions and postconditions that should match to each other while composing, with a few important differences:

- A program is not only associated with a precondition and a postcondition corresponding to the parametrization of the input and output types, but also with **higher order conditions** parameterizing every subtypes (especially arrow types).
- All of these conditions are dependent over "resource variables" that are instantiated by unification during the composition (adding some new constraints). In particular, this allows any program to have a **most general type** that characterizes it; this most general type is computed once and can be reused each time the program is used in a larger project.
- Constraints are not logical propositions but element of an **algebraic structure**. These structures take shape of ordered monoids or ordered semirings. They are reminiscent of abstract interpretation's domains.

The SATIR project has three objectives. The **first** (and main) **objective** is to develop the theory of graded type systems in order to be as expressive as possible, to adapt to existing functional languages and to recover AI's constructions (Galois relations and widening). The **second objective** is to track existing instances of graded type systems and to integrate them in our general framework. The **third objective** is to select a fragment of the system for the purpose of implementation: this fragment need to be inferenceable up-to the use of some widening and calls over an SMT solver (to resolve algebraic constraints).

2 Excellence

2.1 Quality, innovative aspects and credibility of the research (including inter/multidisciplinary aspects)

The SATIR project intends to specify a particular subclass of abstract interpretation that we call graded type systems. Another way to state it is that SATIR project intends to reinject the whole power of abstract interpretation inside a type system by substituting the choice of the domain by the choice of some more refine algebraic structures called grading structures.

Objective 1: Theory Our main objective is to build up a whole theory of graded types in its most general version. This objective can be split in three successive goals. **Goal 1.1** consists in identifying the exact nature of grading structures. **Goal 1.2** aims at transporting abstract interpretation principles to graded type systems. Finally, **Goal 1.3** extends the result in several directions (generalizations to a full fledged language and to intersection type systems).
Objective 2: Applications The second objective tends to stick to real applications. Indeed, we intend to spend an important time looking for existing related work and possible applications. The intent behinds this objective is both to get some inspiration for the general case and to diffuse our results. There are three phases, or goals, regarding the applications. **Goal 2.1** aims at studying and internalizes actual fully formed instances of graded type systems such as those of BLL, D$_\ell$PCF or DFuzz.[1] **Goal 2.2** focuses on qualitative examples and **Goal 2.3** focuses on quantitative ones.
Objective 3: Implementation The third objective is the delivery of an implementation for the end of the three years project. This line of work also takes into account the studies relative to the whole type inference process which form **Goal 3.1**. The implementation itself is the **Goal 3.2**, we are aiming for a simplified functional language where the user only provide the graded structure and heuristics for the widening. The objective also contains **Goal 3.3** which is pure learning over later-phase and future complements such as the use of SMT-solvers.

State of the art.

> The most severe restrictions [of type systems] are on the considered properties.

Methodology.

Novelty, timeliness and quality.

2.2 Clarity and quality of transfer of knowledge/training for the development of the researcher in light of the research objectives

2.3 Quality of the supervision and the hosting arrangements

Outgoing Host Institution

Outgoing Phase Group

Outgoing Phase Supervisor

Return Host Institution

Return Host Group

Return Host Supervisor

[1] See paragraph "State of the art"

2.4 Capacity of the researcher to reach and re-enforce a position of professional maturity in research

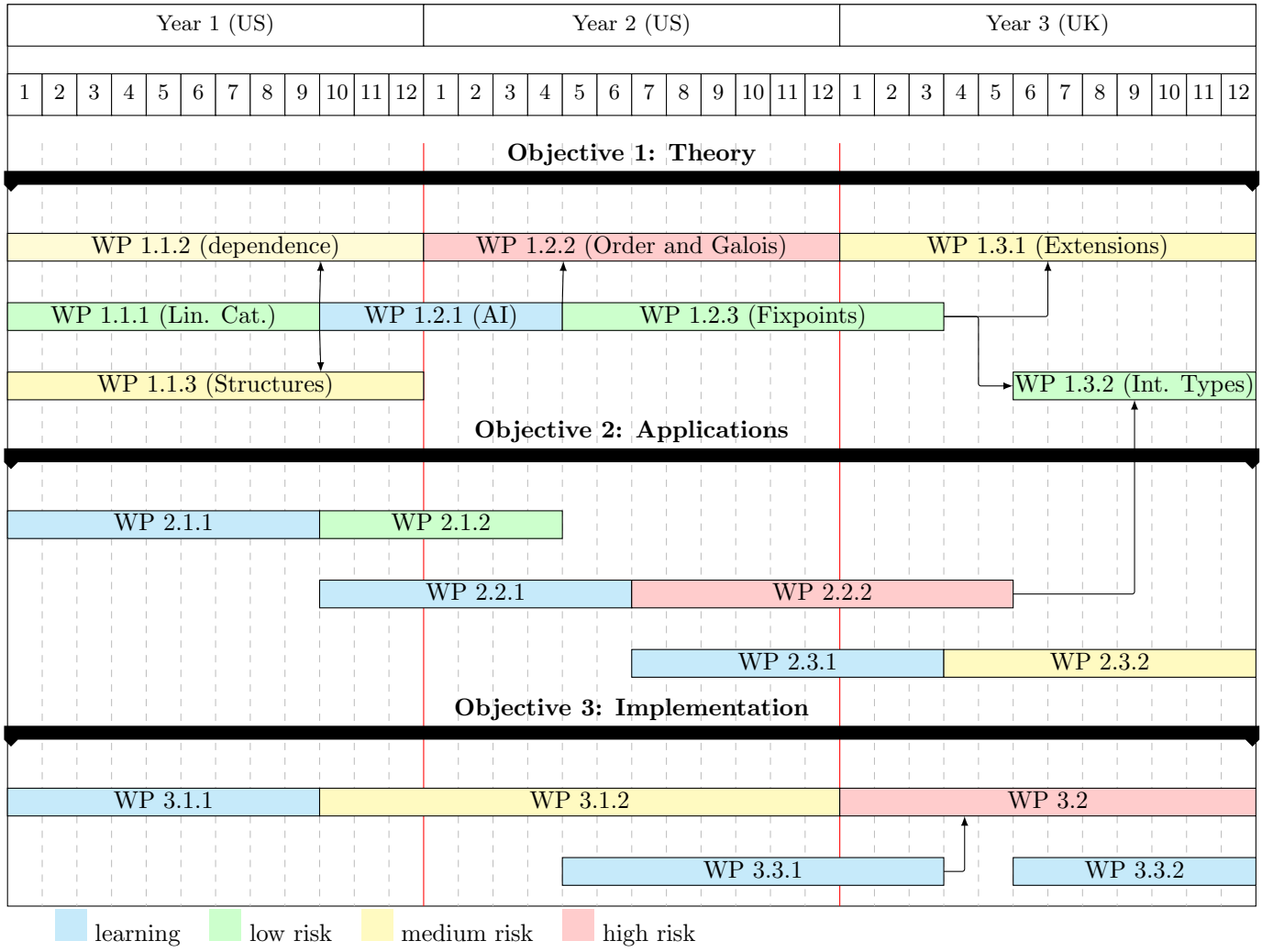
3 Impact

3.1 Enhancing research- and innovation-related human resources, skills, and working conditions to realize the potential of individuals and to provide new career perspectives

3.2 Effectiveness of the proposed measures for communication and results dissemination

4 Implementation

4.1 Overall coherence and effectiveness of the work plan, including appropriateness of the allocation of tasks and resources



The project is divided in three parallel lines. The theory line focuses on developing the main theory behind our project. The Application line has for goal to understand related fields and to apply our result to those fields. The Implementation line is targeting an implementation of our language.

Theory: The main objective follows a sequential pattern.

Goal 1.1 is supposed to be the basis for all the remaining work, it is thus scheduled the first year. Notice, moreover, that no other recherche-focused work package has been planned in parallel (except for the three last months that are assumed to be writing-focused). Goal 1.1 has three different work packages. Work package 1.1.1 is a package that focuses on grading non conventional models of linear logics in order to investigate natural extensions of graded type systems. It is a support package for packages 1.1.2 and 1.1.3, no important direct result are expected here. Work package 1.1.2 has for objective to specify the notion of dependence over grading structures. Work package 1.1.3 has for objective to distinguish a set of independent and interacting layers in the grading structure such as the monad-comonad separation. Early results on these two directions are quite promising and several different peoples are working in this

area; thus expecting a reasonable formalization before the two year deadline (from now) can be considered medium risk.

Goal 1.2 is at the heart of this project and treat the AI-related features of graded type systems. It is planned the second year, but is extended with preliminaries and little remaining work. Work package 1.2.1 is a learning-focused package that aiming at expanding the knowledge of the applicant regarding abstract interpretation. Work package 1.2.2 explores in more details the order that naturally arises from grading structures in order to link the last with domains. We expect this package to end with a notion of Galois relation between grading structures. Work package 1.2.3 is focusing on fixpoints, and in particular on the corresponding notion of widening. This package is considered low risk once the Galois connection has been properly set.

Goal 1.3 consists in optional extensions of the theoretical work. In particular, Work package 1.3.1 is aimed at extending the theory for a λ-calculus to a full fledged language and Work package 1.3.2 is investigating the extension to intersection types.

Application: The application line follows a classic 2-2-2 pattern corresponding to the three applicative objectives, each separated into a learning phase (9 mouths) and an active phase (duration varying with the difficulty level). Notice that this objective is composed only of collaborative material.

Goal 2.1 (work packages 2.1.1 and 2.2) targets the embedding of BLL, D_ℓPCF and DFuzz in our system. It is considered low risk as these languages are well established so that we roughly have to verify that we are taking into consideration each of its components. This objective is mainly aiming at supporting Goal 1.1 with real examples. Moreover, this goal has been scheduled just after one year spent in Focus team (Bologna) that created D_ℓPCF and in the UPen's programming language team that created DFuzz.

Goal 2.2 (work packages 2.2.1 and 2.2.2) targets the embedding of qualitative examples such as the control flow analyzes and[2] the higher order model checking. This objective is classed high risk for several reasons: this will be the first new application and the relation between graded type systems and higher order model checking is not yet fully established. Its scheduling at middle position among applicative goals permit to manage the high risk by potentially extending the deadline. Moreover, the parallelizing of Goal 2.2 and the work package 1.2.3 is important as the fixpoint construction is fundamental in higher order model checking application. Finally, its scheduling permit a potential additional investigation over intersection type systems at the end of the project.

Goal 2.3 (work packages 2.3.1 and 2.3.2) targets the embedding of quantitative examples such as probability or scheduling. This goal is classified medium risk despite their intrinsic difficulties since the applicant has some experience with these cases. Remark that this goal is planed in Birmingham with Dan Ghica that is an expert of the scheduling example (which we believe critical).

Implementation:

Not yet fixed
WP 3.1.1: learning over type inference
WP 3.1.2: working on type inference of graded type systems
WP 3.2: concrete implementation
WP 3.3.1: Learning over SMT solvers
WP 3.3.2: optional learning

Milestones and deliverables: There is three main milestone at the end of every years and three secondary milestone around midyear.

The first secondary milestone (around month 9) do not comes with any delivery.It marks the end of the adaptation period. In particular, WP 2.1 and WP 3.1 will switch from learning phase to active file while WP 1.1 will go to a writing phase. The end of this writing phase is marked by the **Milestone 1** with 2 deliverable articles corresponding to WP 1.1.2 and 1.1.3; at least one of these article should be a collaboration with the host team, corresponding with the ongoing WP 2.1.2.

or col. with Steve over WP 1.1

The second secondary milestone (from month 16 to month 18) is a fuzzy transition with an collaborative article as delivery focused on Objective 2.1 and with early results of WP 3.1.2. **Milestone 2** corresponds to the delivery of two articles. The first is a collaborative article treating WP 1.2.2 and the first advancements over WP 2.2.2. The second is a single-authored article over WP 3.1.2.

The last secondary milestone (from month 27 to month 29) is an other fuzzy transition with a collaborative article as delivery around work packages 1.2.3 and 2.2.2; this milestone will also mark the point where an engineer will be hired to work on the implementation.**Milestone 3** marks the end of the project with the full implementation of WP 3.2 as delivery as well as a collaborative article over WP 2.3.2 together with results from WP 1.3.1.

to discuss

[2] the intersection-free fragment of

4.2 Appropriateness of the management structure and procedures, including quality management and risk management

4.3 Appropriateness of the institutional environment (infrastructure)

4.4 Competences, experience and complementarity of the participating organisations and institutional commitment

5 CV of the Experienced Researcher

Personal Informations

Name Flavien BREUVART

Birthday April 6th, 1988

Family Married

Email breuvart@pps.univ-paris-diderot.fr

Website http://www.pps.univ-paris-diderot.fr/ breuvart/

Position

2015-2017: **Iria post-doctoral positions,** *in Focus Team,* Bologna, Italy

2012-2015: **PhD candidate,** with Antonio Bucciarelli and Michele Pagani, *at PPS, University Paris Diderot,* Paris, France

2012-2015: **Teaching Assistant,** *at UFR Informatique, University Paris Diderot,* Paris, France

Other research experience

2011-2012: **12 months internship,** with Antonio Bucciarelli, *at PPS, University Paris Diderot,* Paris, France

2011: **5 months internship,** with Thomas Ehrhard and Antonio Bucciarelli, *at PPS, University Paris Diderot,* Paris, France

2010: **5 months internship,** with Peter Selinger, *at Dalhousie University,* Halifax, Canada

2009: **2 months internship,** with Achim Jung, *at University of Birmingham,* Birmingham, UK

Grants and Awards

2015: **Iria post-doctoral positions,** *in Focus Team,* Bologna, Italy

2014: **Kleene Award,** for best student paper, *in CSL-LICS,* Vienna, Austria

2012: **Three-Years PhD fellowship,** *from the french "Ministry of Higher Education and Research",* France

Publications

Conference papers

CSL'15: **Breuvart Flavien and Pagani Michele,** *Berlin, Germany*
"Modelling Coeffects in the Relational Semantics of Linear Logic"

CSL-LICS'14 **Breuvart Flavien** *Vienna, Austria*
"On the characterization of models of $\mathcal{H}^*$"

TLCA'13 **Breuvart Flavien** *Eindhoven, Netherlands*
"The resource lambda calculus is short-sighted in its relational model"

Submited

2016 **Breuvart Flavien, Marco Gaboardi, Shin-ya Katsumata and Dominic Orchard,**
"Combining effects and coeffects"

2016 **Breuvart Flavien,** long version
"On the characterization of models of $\mathcal{H}^*$"

non peer-reviewed

2012 **Breuvart Flavien,** arXiv preprint arXiv:1205.4691,
"On the discriminating power of tests in resource lambda-calculus"

Responsibilities

2014: **workshop organization,** Participation at the organization of the international workshop Domain XI, *at PPS, University Paris Diderot,* Paris, France

2013: **seminar organization,** Organization of the temporary working group "Sémantique des calculs classiques", *at PPS, University Paris Diderot,* Paris, France

2012-2015: **Representative of PhD students,** in the Board of directors, *of the laboratory PPS,* Paris, France

Teaching

2015: **advanced Java,** TA for L3 course named POOIG, *at university Paris Diderot,* Paris, France

2014: **advanced Java,** TA for M3 course for linguists named POO3, *at university Paris Diderot,* Paris, France

2014 & 2015: **architecture,** TA for M1 course, *at university Paris Diderot,* Paris, France

2014: **"UE libre" vulgarization,** In charge of the "UE libre" (optional short course) on scientific vulgarization in M1, *at university Paris Diderot,* Paris, France

2013 & 2014: **mid-beginner Java,** TA for L1 course named TO2, *at university Paris Diderot,* Paris, France

2012: **Java projects,** Supervising Java projects in L2 (more exactly for the "classe préparatoire intégrée de l'EIDD"), *at university Paris Diderot,* Paris, France

2008-2012: **vulgarization,** occasional (5-10 times per years) 1 hour of vulgarization on mathematics and computer sciences for middle school students, *with the association* PRACCIS*,* Cachan, France

2008-2009: **vulgarization of game theory,** 1h/week of vulgarization of game theory to high school students, *with the association* Sciences Ouvertes*,* Drancy, France

Communications

10/09/2015 **CSL'2015,** Berlin, Germany
"Modelling Coeffects in Rel"

03/09/2014 **seminar of logic team,** of LAMA, *University of Savoie,* Chambéry, France
"On the characterization of models of $\mathcal{H}^*$"

17/07/2014 **CSL-LICS,** Vienna, Austria
"On the characterization of models of $\mathcal{H}^*$"

04/04/2014 **seminar of team LCR,** in LIPN, *Univerisity Paris Nort,* Paris, France
"De la charactérisation des modèles de $\mathcal{H}^*$"

09/04/2014 **LIAFA-PPS' PHD seminar,** *University Paris Diderot,* Paris,France
"Basic notions on weighted models"

14/03/2014 **"Theory seminar",** School of Computer Science, *University of Birmingham,* Birmingham, UK
"A bridge between semirings"

03/12/2013 **BLL'workshop,** Fontainebleau, France
"A bridge between semirings"

27/06/2013 **TLCA,** Eindhoven, Netherlands
"The resource lambda calculus is short-sighted in its relational model"

15/02/2013 **GEOCAL Meeting,** Lyon, France
"Une bien étrange caractérisation de l'adéquation complète dans les domaines de Scott"

24/10/2012 **LIAFA-PPS' PHD seminar,** *University Paris Diderot,* Paris,France
"Petites histoires entre modèles et syntaxes"

12/09/2012 **PPS' anual meating,** *University Paris Diderot,* Trouville,France
"Petites histoires entre modèles et syntaxes"

02/09/2011 **LIAFA-PPS' PHD seminar,** *University Paris Diderot,* Paris,France
"Et si les PPSiens se mettaient à faire de l'algèbre?!!"

Education

2010-2011: **PHD thesis in Paris Diderot,** *Paris, France*

2010-2011: **M2 in MPRI,** *Cachan, France*

2009-1010: **M1 in MPRI,** Cachan, France

2008-2009: **L3 in computer sciences and in mathematics,** *at ENS Cachan,* Cachan, France

2008: **admission at ENS Cachan,** Cachan, France

2006-2008: **CPGE MPSI and MP*,** at Lycée Fénélon, Paris 5$^{\text{iem}}$, France

2006: **French baccalaureate,** "mention bien", *at Lycée W.A.Mozart,,* Le Blanc Mesnil, France

Associative activities

2010-2011: **president of *PRACCIS* project,** in *SONO*, Cachan, France

2009-2010: **active member,** of *BDE ENS Cachan*, *with high implication in the clubs* BDA, Mediatek, PRACCIS Cachan,France

2007-2008: **active member** of *Sciences Ouvertes* Drancy, France

2005-2007: **board of directors** of the association *MJC Daniel André* Drancy, France

6 Capacities of the Participating Organisations

7 Ethical Aspects

There are no ethical sensitive issues associated with the subject of this proposal.

8 Letters of Commitment of Partner Organisations