Differential Logical Relations Joint work with Francesco Gavazzo and Akira Yoshimizu

Ugo Dal Lago



IFIP WG 2.2 Annual Meeting, Vienna, September 23rd 2019





European Research Council

Established by the European Commission

This work has been supported by the ERC CoG DIAPASoN, GA 818616.

Comparing Interacting Programs



Comparing Interacting Programs



► Let $\equiv \subseteq \Lambda \times \Lambda$ be a notion of equivalence. What are the minimal requirements we should put on \equiv ?

- ► Let $\equiv \subseteq \Lambda \times \Lambda$ be a notion of equivalence. What are the minimal requirements we should put on \equiv ?
- Adequacy: for every M, N,

$$M \equiv N \implies Obs(M) = Obs(N)$$

where $Obs : \Lambda \to \mathbb{X}$ is the observation function.

► Let $\equiv \subseteq \Lambda \times \Lambda$ be a notion of equivalence. What are the minimal requirements we should put on \equiv ?

• Adequacy: for every
$$M, N$$
,

$$M \equiv N \implies Obs(M) = Obs(N)$$

where $Obs : \Lambda \to \mathbb{X}$ is the observation function.

Congruence: for every M, N, C,

$$M\equiv N \implies C[M]\equiv C[N].$$

- ► Let $\equiv \subseteq \Lambda \times \Lambda$ be a notion of equivalence. What are the minimal requirements we should put on \equiv ?
- Adequacy: for every M, N,

$$M \equiv N \implies Obs(M) = Obs(N)$$

where $Obs : \Lambda \to \mathbb{X}$ is the observation function.

Congruence: for every M, N, C,

$$M\equiv N \implies C[M]\equiv C[N].$$

• Examples, the realm of λ -calculus:

- ► The largest adequate congruence, **context equivalence** [Morris1968].
- ▶ Logical relations [Plotkin1973].
- Applicative [Abramsky1990] or open bisimilarity.

- ► Let $\equiv \subseteq \Lambda \times \Lambda$ be a notion of equivalence. What are the minimal requirements we should put on \equiv ?
- Adequacy: for every M, N,

$$M\equiv N \implies Obs(M)=Obs(N)$$

where $Obs : \Lambda \to \mathbb{X}$ is the observation function.

Congruence: for every M, N, C,

$$M\equiv N \implies C[M]\equiv C[N].$$

• Examples, the realm of λ -calculus:

- The largest adequate congruence, context equivalence [Morris1968].
- **Logical relations** [Plotkin1973].
- Applicative [Abramsky1990] or open bisimilarity.
- ▶ What if X is a metric space?

$$\blacktriangleright \text{ Let } \delta : \Lambda \times \Lambda \to \mathbb{R}.$$

- $\blacktriangleright \text{ Let } \delta : \Lambda \times \Lambda \to \mathbb{R}.$
- Adequacy: for every M, N,

 $\delta(M,N) \geq \delta_{\mathbb{X}}(Obs(M),Obs(N))$

- $\blacktriangleright \text{ Let } \delta : \Lambda \times \Lambda \to \mathbb{R}.$
- Adequacy: for every M, N,

 $\delta(M,N) \geq \delta_{\mathbb{X}}(Obs(M),Obs(N))$

▶ Non-Expansiveness: for every M, N, C,

 $\delta(M,N) \geq \delta(C[M],C[N])$

- $\blacktriangleright \text{ Let } \delta : \Lambda \times \Lambda \to \mathbb{R}.$
- Adequacy: for every M, N,

 $\delta(M,N) \geq \delta_{\mathbb{X}}(Obs(M),Obs(N))$

▶ Non-Expansiveness: for every M, N, C,

 $\delta(M,N) \geq \delta(C[M],C[N])$

▶ In *probabilistic* computation, one is naturally lead to observe a quantitative property, and X is simply R.

- $\blacktriangleright \text{ Let } \delta : \Lambda \times \Lambda \to \mathbb{R}.$
- Adequacy: for every M, N,

 $\delta(M,N) \geq \delta_{\mathbb{X}}(Obs(M),Obs(N))$

▶ Non-Expansiveness: for every M, N, C,

 $\delta(M,N) \geq \delta(C[M],C[N])$

- ▶ In *probabilistic* computation, one is naturally lead to observe a quantitative property, and X is simply R.
- ▶ But even when computation is *deterministic*, one could well work with X = R when real numbers are part of the underlying language [ReedPierce2010,AGHKC2017].

A Survey Of Techniques for Approximate Computing

Sparsh Mittal, Oak Ridge National Laboratory

Approximate computing trades off computation quality with the effort expended and as rising performance demands confront with plateauing resource budgets, approximate computing has become, not merely attractive, but even imperative. In this paper, we present a survey of techniques for approximate computing (AC). We discuss strategies for finding approximable program portions and monitoring output quality, techniques for using AC in different processing units (e.g., CPU, GPU and FPGA), processor components, memory technologies etc., and programming frameworks for AC. We classify these techniques has do as everal kay characteristics to emphasize their similarities and differences. The aim of this paper is to provide insights to researchers into working of AC techniques and inspire more efforts in this area to make AC the mainstream computing approach in future systems.

Categories and Subject Descriptors: [General and reference]: Surveys and overviews; [Hardware]: Power and energy; [Computer systems organization]: Processors and memory architectures

General Terms: Design, Performance

Additional Key Words and Phrases: Review, classification, approximate computing technique (ACT), approximate storage, quality configurability, CPU, GPU, FPGA, neural networks

ACM Reference Format:

S. Mittal, "A Survey Of Techniques for Approximate Computing", 20xx. ACM Comput. Surv. a, b, Article 1 (2016), 34 pages. DOI: http://x.doi.org/10.1145/000000.0000000

1. INTRODUCTION

As large-scale applications such as scientific computing, social media and financial analysis gain prominence, the computational and storage demands of modern systems

$$M_{ID} \equiv \lambda x.x$$
 $M_{SIN} \equiv \lambda x.\sin x$





 $\delta(M_{ID}, M_{SIN}) = +\infty$



$\delta(M_{ID}, M_{SIN}) = +\infty$

What if the *environment* feeds the function with values close to 0, only?

A Semantics for Approximate Program Transformations

Edwin Westbrook and Swarat Chaudhuri Department of Computer Science, Rice University Houston, TX 77005 Email: {emv4.swarat}@rice.edu

Abstract

An approximate program transformation is a transformation that can change the semantics of a program within a specified empirical error bound. Such transformations have wide applications: they can decrease computation time, power consumption, and memory usage, and can, in some cases, allow implementations of incomputable operations. Correctness proofs of approximate program transformations are by definition quantitative (Informately, unlike with standard program transformations, there is as of yet no modular way to prove correctness of an approximate transformation itself. Error bounds must be proved for each transformed program individually, and must be reproved each time a program is modified or a different set of approximations are applied. as floating-point numbers, lossy compression, and approximation algorithms for NP-hard problems. Such techniques are often used to trade off accuracy of the result for **reduced resource usage**, for resources such as computation time, power, and memory. In addition, some approximation techniques are also used to ensure **computability**. For example, true representations of real numbers (e.g., [**T**], **II**]), require some operations, such as comparison, to be incomputable; floating-point comparison, in contrast, is efficiently decidable on modern computers.

Recently, there has been a growing interest in language-based approximations, where approximate program transformations are performed by the programming language environment [21], [12], [19], [18], [4], [3], [16]. Such approaches allow the user to give an exact program as a specification, and then apply some

A Toy Language

$\tau, \rho ::= REAL \mid \tau \to \rho \mid \tau \times \rho.$

A Toy Language

Types $\tau, \rho ::= REAL \mid \tau \to \rho \mid \tau \times \rho.$

Typing Rules				
$x:\tau\in\Gamma$		f_n	$\in \mathcal{F}_n$	$\Gamma, x:\tau \vdash M:\rho$
$\overline{\Gamma \vdash x : \tau}$	$\overline{\Gamma \vdash r: REAL}$	$\overline{\Gamma \vdash f_n : RE}$	$EAL^n \to REAL$	$\overline{\Gamma \vdash \lambda x.M: \tau \to \rho}$
$\Gamma \vdash M: \tau \to \rho \exists$	$\Gamma \vdash N : \tau \qquad \Gamma \vdash M :$	$\tau \Gamma \vdash N : \rho$		
$\Gamma \vdash MN$:	ρ $\Gamma \vdash \langle M$	$\langle N \rangle : \tau \times \rho$	$\Gamma \vdash \pi_1 : \tau \times \rho$ -	$\overline{\to \tau}$ $\overline{\Gamma \vdash \pi_2 : \tau \times \rho \to \rho}$
$\frac{\Gamma \vdash M: \tau \Gamma \vdash N: \tau}{\Gamma \vdash \texttt{iflz} \; M \; \texttt{else} \; N: REAL \to \tau}$			$\frac{\Gamma \vdash M: \tau \to \tau \Gamma \vdash N: \tau}{\Gamma \vdash \operatorname{iter} M \text{ base } N: REAL \to \tau}$	

A Toy Language

Types $\tau, \rho ::= REAL \mid \tau \to \rho \mid \tau \times \rho.$

 $\begin{array}{c} \textbf{Denotational Semantics} \\ \llbracket REAL \rrbracket = \mathbb{R}; & \llbracket \tau \to \rho \rrbracket = \llbracket \tau \rrbracket \to \llbracket \rho \rrbracket; & \llbracket \tau \times \rho \rrbracket = \llbracket \tau \rrbracket \times \llbracket \rho \rrbracket. \end{array}$

$$\begin{array}{ll} \textbf{Distance Spaces} \\ (\!(\textit{REAL}) = \mathbb{R}_{\geq 0}^{\infty}; & (\!(\tau \to \rho) \!) = [\![\tau]\!] \times (\!(\tau) \!) \to (\!(\rho)); & (\!(\tau \times \rho) \!) = (\!(\tau)\!) \times (\!(\rho)\!) \end{array}$$



$$\begin{array}{ll} \text{Distance Spaces} \\ (\textit{REAL}) = \mathbb{R}_{\geq 0}^{\infty}; & (\tau \to \rho) = [\![\tau]\!] \times (\![\tau]\!] \to (\![\rho]\!]; & (\tau \times \rho) = (\![\tau]\!] \times (\![\rho]\!] \end{array}$$

DLRs as Ternary Relations

$$\begin{split} \delta_{REAL}(M,r,N) &\Leftrightarrow |NF(M) - NF(N)| \leq r; \\ \delta_{\tau \times \rho}(M,(d_1,d_2),N) &\Leftrightarrow \delta_{\tau}(\pi_1 M, d_1,\pi_1 N) \wedge \delta_{\rho}(\pi_2 M, d_2,\pi_2 N) \\ \delta_{\tau \to \rho}(M,d,N) &\Leftrightarrow (\forall V \in CV(\tau). \; \forall x \in (\!\!(\tau)\!\!). \; \forall W \in CV(\tau). \\ \delta_{\tau}(V,x,W) &\Rightarrow \delta_{\rho}(MV,d([\![V]\!],x),NW) \wedge \delta_{\rho}(MW,d([\![V]\!],x),NV)). \end{split}$$

$$\begin{array}{ll} \textbf{Distance Spaces} \\ (\!(\textit{REAL}\!) = \mathbb{R}_{\geq 0}^{\infty}; & (\!(\tau \to \rho)\!) = [\![\tau]\!] \times (\!(\tau)\!) \to (\!(\rho)\!); & (\!(\tau \times \rho)\!) = (\!(\tau)\!) \times (\!(\rho)\!) \end{array}$$

$\begin{array}{l} \textbf{DLRs as Ternary Relations} \\ \delta_{REAL}(M,r,N) \Leftrightarrow |NF(M) - NF(N)| \leq r; \\ \delta_{\tau \times \rho}(M,(d_1,d_2),N) \Leftrightarrow \delta_{\tau}(\pi_1 M,d_1,\pi_1 N) \wedge \delta_{\rho}(\pi_2 M,d_2,\pi_2 N) \\ \delta_{\tau \to \rho}(M,d,N) \Leftrightarrow (\forall V \in CV(\tau). \ \forall x \in (\![\tau]\!]. \ \forall W \in CV(\tau). \\ \delta_{\tau}(V,x,W) \Rightarrow \delta_{\rho}(MV,d([\![V]\!],x),NW) \wedge \delta_{\rho}(MW,d([\![V]\!],x),NV)). \end{array}$

Theorem (Fundamental Lemma, Version I) For every $\vdash M : \tau$, there is $d \in (|\tau|)$ such that $\delta_{\tau}(M, d, M)$.

▶ Why not null distances?

▶ The distance between a program *M* and itself is null, isn't it?

• Why not null distances?

- ▶ The distance between a program *M* and itself is null, isn't it?
- ▶ In fact, this is true only at ground types.
- Example: the distance between M_{ID} and itself is something like $\lambda\langle x, \varepsilon \rangle.\varepsilon$.

• Why not null distances?

- ▶ The distance between a program *M* and itself is null, isn't it?
- ▶ In fact, this is true only at ground types.
- Example: the distance between M_{ID} and itself is something like $\lambda\langle x, \varepsilon \rangle.\varepsilon$.

Too weak?

- The distance d is arbitrary, and can even be infinite.
- ▶ In ordinary logical relations, the FL enables compositional reasoning, when applied to the environment, the context.

• Why not null distances?

- ▶ The distance between a program *M* and itself is null, isn't it?
- ▶ In fact, this is true only at ground types.
- Example: the distance between M_{ID} and itself is something like $\lambda\langle x, \varepsilon \rangle.\varepsilon$.

▶ Too weak?

- The distance d is arbitrary, and can even be infinite.
- ▶ In ordinary logical relations, the FL enables compositional reasoning, when applied to the environment, the context.
- ▶ The same here:

• Why not null distances?

- ▶ The distance between a program *M* and itself is null, isn't it?
- ▶ In fact, this is true only at ground types.
- Example: the distance between M_{ID} and itself is something like $\lambda\langle x, \varepsilon \rangle.\varepsilon$.

Too weak?

- The distance d is arbitrary, and can even be infinite.
- In ordinary logical relations, the FL enables compositional reasoning, when applied to the environment, the context.

▶ The same here:

 $\begin{array}{c} C:\tau \rightarrow REAL\\ M,N:\tau \end{array}$

• Why not null distances?

- ▶ The distance between a program *M* and itself is null, isn't it?
- ▶ In fact, this is true only at ground types.
- Example: the distance between M_{ID} and itself is something like $\lambda\langle x, \varepsilon \rangle.\varepsilon$.

▶ Too weak?

- The distance d is arbitrary, and can even be infinite.
- In ordinary logical relations, the FL enables compositional reasoning, when applied to the environment, the context.
- The same here:

$$C: \tau \to REAL \qquad (C, d, C) \in \delta_{\tau \to REAL}$$
$$M, N: \tau \qquad (M, e, N) \in \delta_{\tau}$$

• Why not null distances?

- ▶ The distance between a program *M* and itself is null, isn't it?
- ▶ In fact, this is true only at ground types.
- Example: the distance between M_{ID} and itself is something like $\lambda\langle x, \varepsilon \rangle.\varepsilon$.

Too weak?

- The distance d is arbitrary, and can even be infinite.
- ▶ In ordinary logical relations, the FL enables compositional reasoning, when applied to the environment, the context.

The same here:

$$C: \tau \to REAL \qquad (C, d, C) \in \delta_{\tau \to REAL}$$
$$M, N: \tau \qquad (M, e, N) \in \delta_{\tau}$$
$$\Downarrow \qquad (C[M], d(\llbracket M \rrbracket, e), C[N]) \in \delta_{REAL}$$

Claim

$$\delta_{REAL \to REAL}(M_{ID}, \lambda \langle x, y \rangle . y + |x - \sin x|, M_{SIN})$$

Claim

$$\delta_{REAL \to REAL}(M_{ID}, \lambda \langle x, y \rangle . y + |x - \sin x|, M_{SIN})$$

Proof.

Consider any pairs of real numbers $r, s \in \mathbb{R}$ such that $|r-s| \leq \varepsilon$, where $\varepsilon \in \mathbb{R}_{\geq 0}^{\infty}$. We have that:

$$\begin{split} |\sin r - s| &= |\sin r - r + r - s| \le |\sin r - r| + |r - s| \\ &\le |\sin r - r| + \varepsilon = f(r, \varepsilon) \\ |\sin s - r| &= |\sin s - \sin r + \sin r - r| \\ &\le |\sin s - \sin r| + |\sin r - r| \le |s - r| + |\sin r - r| \\ &\le \varepsilon + |\sin r - r| = f(r, \varepsilon). \end{split}$$

where $f = \lambda \langle x, y \rangle \cdot y + |x - \sin x|$.

Now, consider, e.g., the context $C = (\lambda x. x(x\theta))[\cdot]$.

- ► Now, consider, e.g., the context $C = (\lambda x.x(x\theta))[\cdot]$.
- C can be seen as a term having type $\tau = (REALS \rightarrow REALS) \rightarrow REALS$. A self-distance d for C can thus be defined as an element of

$$(\tau) = [[REALS \to REALS]] \times ([REALS \to REALS]) \to [[REALS]].$$

namely $F = \lambda \langle g, h \rangle .h(g(\theta), h(\theta, 0)).$

- ► Now, consider, e.g., the context $C = (\lambda x.x(x\theta))[\cdot]$.
- C can be seen as a term having type $\tau = (REALS \rightarrow REALS) \rightarrow REALS$. A self-distance d for C can thus be defined as an element of

 $(\!(\tau)\!) = [\![REALS \to REALS]\!] \times (\![REALS \to REALS]\!) \to [\![REALS]\!].$

namely $F = \lambda \langle g, h \rangle .h(g(\theta), h(\theta, 0)).$

▶ This allows for **compositional reasoning** about program distances: the overall impact of replacing M_{SIN} by M_{ID} can be evaluated by computing $F(\llbracket M_{ID} \rrbracket, f)$ or $F(\llbracket M_{SIN} \rrbracket, f)$.

- ► Now, consider, e.g., the context $C = (\lambda x.x(x\theta))[\cdot]$.
- C can be seen as a term having type $\tau = (REALS \rightarrow REALS) \rightarrow REALS$. A self-distance d for C can thus be defined as an element of

 $(\tau) = [\![REALS \to REALS]\!] \times (\![REALS \to REALS]\!) \to [\![REALS]\!].$

namely $F = \lambda \langle g, h \rangle .h(g(\theta), h(\theta, 0)).$

- ▶ This allows for **compositional reasoning** about program distances: the overall impact of replacing M_{SIN} by M_{ID} can be evaluated by computing $F(\llbracket M_{ID} \rrbracket, f)$ or $F(\llbracket M_{SIN} \rrbracket, f)$.
- Of course the context C needs to be taken into account, but **once and for all**: the functional F can be built without knowing either M_{SIN} or M_{ID} .

Hereditarily Null Distances $(REAL)^{0} = \{0\} \qquad (\tau \times \rho)^{0} = (\tau)^{0} \times (\rho)^{0}$ $(\tau \to \rho)^{0} = \{f \mid \forall x \in [[\tau]]. \forall y \in (\tau)^{0}. f(x, y) \in (\rho)^{0}\}$



Hereditarily Null Distances $(REAL)^{0} = \{0\} \qquad (\tau \times \rho)^{0} = (\tau)^{0} \times (\rho)^{0}$ $(\tau \to \rho)^{0} = \{f \mid \forall x \in [[\tau]] . \forall y \in (\tau)^{0} . f(x, y) \in (\rho)^{0}\}$

Hereditarily Finite Distances

 $(REAL)^{<\infty} = \mathbb{R}_{\geq 0}; \qquad (\tau \times \rho)^{<\infty} = (\tau)^{<\infty} \times (\rho)^{<\infty};$

 $(\!(\tau \to \rho)\!)^{<\infty} = \{ f \in (\!(\tau \to \rho)\!) \mid \forall x \in [\![\tau]\!] . \forall t \in (\!(\tau)\!)^{<\infty} . f(x,t) \in (\!(\rho)\!)^{<\infty} \}.$

Hereditarily Null Distances $(REAL)^{0} = \{0\} \qquad (\tau \times \rho)^{0} = (\tau)^{0} \times (\rho)^{0}$ $(\tau \to \rho)^{0} = \{f \mid \forall x \in [[\tau]] . \forall y \in (\tau)^{0} . f(x, y) \in (\rho)^{0}\}$

Hereditarily Finite Distances

 $(REAL)^{<\infty} = \mathbb{R}_{\geq 0}; \qquad (\tau \times \rho)^{<\infty} = (\tau)^{<\infty} \times (\rho)^{<\infty};$

$$(\! | \tau \to \rho)\! |^{<\infty} = \{ f \in (\! | \tau \to \rho)\! \mid \forall x \in [\! [\tau]\!]. \forall t \in (\! | \tau)\! |^{<\infty}. f(x,t) \in (\! | \rho)\! |^{<\infty} \}.$$

Lemma

Whenever $\vdash M, N : \tau, M$ is logically related to N iff $\delta_{\tau}(M, d, N)$ where $d \in (|\tau|)^0$.

Hereditarily Null Distances $(REAL)^{0} = \{0\} \qquad (\tau \times \rho)^{0} = (\tau)^{0} \times (\rho)^{0}$ $(\tau \to \rho)^{0} = \{f \mid \forall x \in [[\tau]] . \forall y \in (\tau)^{0} . f(x, y) \in (\rho)^{0}\}$

Hereditarily Finite Distances

 $(REAL)^{<\infty} = \mathbb{R}_{\geq 0}; \qquad (\tau \times \rho)^{<\infty} = (\tau)^{<\infty} \times (\rho)^{<\infty};$

$$(\!\! | \tau \to \rho)\!\! |^{<\infty} = \{ f \in (\!\! | \tau \to \rho)\!\! | \ \forall x \in [\!\! | \tau]\!\!]. \forall t \in (\!\! | \tau)\!\! |^{<\infty}. f(x,t) \in (\!\! | \rho)\!\! |^{<\infty} \}.$$

Lemma

Whenever $\vdash M, N : \tau, M$ is logically related to N iff $\delta_{\tau}(M, d, N)$ where $d \in (|\tau|)^0$.

Theorem (Fundamental Lemma, Version II) For every $\vdash M : \tau$, there is $d \in (|\tau|)^{<\infty}$ such that $\delta_{\tau}(M, d, M)$.

Conclusions

Other Interesting Results

- Differential logical relations are examples of generalized metric domains, which (contrarily to metric spaces) form a *cartesian closed category*.
- Not only logical relations, but also metric logical relations can be seen as special kinds of distances.

Conclusions

Other Interesting Results

- Differential logical relations are examples of generalized metric domains, which (contrarily to metric spaces) form a *cartesian cased category*.
- Not only logical relations, but also metric logical relations

$$\begin{split} \delta(x,0,y) \Rightarrow x &= y \\ \delta(x,d,y) \Rightarrow \delta(y,d,x) \\ \delta(x,d,y) \wedge \delta(y,e,y) \wedge \delta(y,f,z) \Rightarrow \delta(x,d+e+f,z) \end{split}$$

Conclusions

Other Interesting Results

- Differential logical relations are examples of generalized metric domains, which (contrarily to metric spaces) form a *cartesian closed category*.
- Not only logical relations, but also metric logical relations can be seen as special kinds of distances.

Further Work

- Capture more expressive constructs and calculi (e.g. recursion by way of step-indexing).
- ► Higher-order continuity?
- ▶ Abstracting functional distances by way of step functions.
- ▶ Make behavioral metrics *context-dependent*.

Thank You!

Questions?