On Envelopes and Backward Approximations

Eike Neumann

Max Planck Institute for Software Systems, Saarbrücken, EU

CCC 2021, Birmingham, UK 24 September 2021

Acknowledgements

This work is heavily inspired by discussions with Franz Brausse, Pieter Collins, Michal Konečný, Norbert Müller, Sewon Park, Florian Steinberg, and Martin Ziegler during my secondments within the CID project at KAIST in 2017 and 2019.

Folklore observations surrounding comparison of real numbers for inequality.

$$\frac{1}{\sqrt{2\pi}} \int_0^\infty x^{\sqrt{1729}} e^{-x} dx \stackrel{?}{\leq} \sqrt{1729}^{\sqrt{1729} + \frac{1}{2}} e^{-\sqrt{1729}} e^{\frac{1}{12\sqrt{1729}}}$$

Let's compute some functions!

- **②** Suppose we have implemented algorithms for computing functions $f: \mathbb{R} \to \mathbb{R}$ and $g: \mathbb{R} \to \mathbb{R}$.
- 2 Let's use these to compute a new function!
- Why not take

$$h(x) = \begin{cases} f(x) & \text{if } x \ge 0, \\ g(x) & \text{if } x < 0. \end{cases}$$

4 Here's an algorithm that computes it:

```
if (x \ge 0) then:

return f(x);

else:

return g(x);
```

Let's compute some functions!

- **②** Suppose we have implemented algorithms for computing functions $f: \mathbb{R} \to \mathbb{R}$ and $g: \mathbb{R} \to \mathbb{R}$.
- 2 Let's use these to compute a new function!
- Why not take

$$h(x) = \begin{cases} f(x) & \text{if } x \ge 0, \\ g(x) & \text{if } x < 0. \end{cases}$$

4 Here's an algorithm that computes it:

```
Undecidable!

if (x \ge 0) then:

return f(x);

else:

return g(x);
```

First solution: replace \leq : $\mathbb{R}^2 \rightarrow \{\text{True}, \text{False}\}\$ with

$$\leq$$
: $\mathbb{R}^2 \to \{\text{True}, \text{False}, \bot\}, (x \leq y) = \begin{cases} \text{True} & \text{if } x < y \\ \text{False} & \text{if } x > y \\ \bot & \text{if } x = y. \end{cases}$

Equivalently,

$$\leq : \mathbb{R}^2 \to \mathcal{K}(\{\mathsf{True},\mathsf{False}\}), \ (x \leq y) = \begin{cases} \{\mathsf{True}\} & \text{if } x > y \\ \{\mathsf{False}\} & \text{if } x < y \\ \{\mathsf{True},\mathsf{False}\} & \text{if } x = y. \end{cases}$$

```
if (x \ge 0) then:

return f(x);

else:

return g(x);
```

```
x = 0
if (x \ge 0) then:
return f(x);
else:
return g(x);
```

```
x = 0
if (0 \ge 0) then:
return f(x);
else:
return g(x);
```

```
x = 0
if (\{True, False\}) then:
return f(x);
else:
return g(x);
```

```
x = 0
if (\{True, False\}) then:
return f(0);
else:
return g(x);
```

```
x = 0
if (\{True, False\}) then:
return f(0);
else:
return g(0);
```

```
 \begin{array}{c} x=0 \\ \text{if } \left( \{ \textit{True}, \textit{False} \} \right) \text{ then:} \\ & \quad \text{return } f(0); \\ \text{else:} \\ & \quad \text{return } g(0); \\ \\ \rightsquigarrow \text{value} = \left\{ f(0), g(0) \right\} \in \mathcal{K} \left( \{ \text{True}, \text{False} \} \right) \end{array}
```

Second solution: replace \leq : $\mathbb{R}^2 \rightarrow \{\text{True}, \text{False}\}\$ with

$$^{\dagger} \leq : \mathbb{R}^2 \times \mathbb{Q}_{>0} \leadsto \{\mathsf{True}, \mathsf{False}\}, \ \left(x^{\dagger} \leq_{\delta} y\right) = \begin{cases} \{\mathsf{True}\} & \text{if } x \geq y + 2\delta \\ \{\mathsf{False}\} & \text{if } x \leq y - 2\delta \\ \{\mathsf{True}, \mathsf{False}\} & \text{if } |x - y| < 2\delta. \end{cases}$$

```
if (x^{\dagger} \ge_{\delta} 0) then:

return f(x);

else:

return g(x);
```

```
\eta \in (-2\delta, 2\delta)

if (x^{\dagger} \ge_{\delta} 0) then:

return f(x);

else:

return g(x);
```

```
\eta \in (-2\delta, 2\delta)

if (\eta^{\dagger} \ge_{\delta} 0) then:

return f(x);

else:

return g(x);
```

```
Fix \delta > 0. Run the algorithm using lower semantics: \eta \in (-2\delta, 2\delta) if (\mathit{True}) then: Non-deterministic choice. return f(x); else: return g(x);
```

```
Fix \delta > 0. Run the algorithm using lower semantics: \eta \in (-2\delta, 2\delta) if (\mathit{True}) then: Non-deterministic choice. return f(\eta); else: return g(x);
```

```
\eta \in (-2\delta, 2\delta)

if (True) then: Non-deterministic choice.

return f(\eta);

else:

return g(x);

\rightarrow value \in \{f(\eta), g(\eta)\}
```

```
\begin{split} \eta &\in (-2\delta, 2\delta) \\ &\quad \text{if } (\textit{True}) \text{ then: Non-deterministic choice.} \\ &\quad \textit{return } f(\eta); \\ &\quad \text{else:} \\ &\quad \textit{return } g(x); \\ &\quad \rightsquigarrow \textit{value } \in \{f(\eta), g(\eta)\} \\ \Rightarrow \varepsilon\text{-close to } h(\eta) \text{ if } \delta \in \min \{\omega_f(0, \varepsilon/2), \omega_g(0, \varepsilon/2)\}. \end{split}
```

```
Fix \delta > 0. Run the algorithm using lower semantics:
          \eta \in (-2\delta, 2\delta)
                  if (True) then: Non-deterministic choice.
                        return f(\eta);
                  else:
                        return g(x);
           \rightsquigarrow value \in \{f(\eta), g(\eta)\}
           \Rightarrow \varepsilon-close to h(\eta) if \delta \in \min \{ \omega_f(0, \varepsilon/2), \omega_g(0, \varepsilon/2) \}.
                                                         moduli of continuity
```

Summary

Naïve semantics:

if
$$(x \ge 0)$$
 then:
return $f(x)$;

else:

return
$$g(x)$$
;

Obviously correct. Easiest to understand. Impossible to execute on a machine. Upper semantics:

if
$$(x \ge 0)$$
 then:
return $f(x)$;

else:

return
$$g(x)$$
;

$$f(0)=g(0).$$

Slightly harder to understand. Can be executed. Quite inefficient. Lower semantics:

if
$$(x^{\dagger} \ge_{\delta} 0)$$
 then:
return $f(x)$;
else:

return
$$g(x)$$
;

$$\delta \in \min \{ \omega_f(0, \varepsilon/2), \\ \omega_g(0, \varepsilon/2) \}.$$

Hardest to understand. Correctness requires "hard analysis". Can be executed efficiently.

Let's compute more functions!

- Goal: relate "upper" and "lower" semantics of "arbitrary" (naïve) programs.
- For now, focus on chains of function compositions:

$$f_n \circ \cdots \circ f_1$$

where $f_1, ..., f_n$ are arbitrary (!) functions between effective metric spaces.

Backward Approximations

Let $f: X \to Y$ be an arbitrary (!) function between effective metric spaces. Let

$$^{\dagger}f: X \times \mathbb{Q}_{>0} \leadsto Y, \,^{\dagger}f(x,\delta) = \{f(\widetilde{x}) \mid \widetilde{x} \in B(x,\delta)\}$$

be its backward approximation.

- Very common relaxation: equation solving, matrix diagonalisation, backwards stable algorithms...
- Useful for computing functions that do depend continuously on the input.
- \bullet † f computable if f has a computable left inverse.

Backward Approximations

Idea: instead of computing

$$f_n \circ \cdots \circ f_1(x)$$

compute

$$^{\dagger}f_n(\cdot,\delta_n)\circ\cdots\circ^{\dagger}f_1(\cdot,\delta_1)$$

for "sufficiently small" δ_i 's and hope that the result is close to the true result.

Backward Approximations

Idea: instead of computing

$$f_n \circ \cdots \circ f_1(x)$$

compute

$$^{\dagger}f_n(\cdot,\delta_n)\circ\cdots\circ^{\dagger}f_1(\cdot,\delta_1)$$

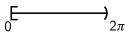
for "sufficiently small" δ_i 's and hope that the result is close to the true result.

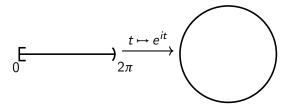
Main Question

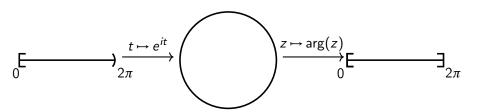
Let $x \in X$. Under which assumptions on f_1, \ldots, f_n is it possible to find for all $\varepsilon > 0$ numbers $\delta_1 > 0, \ldots, \delta_n > 0$ such that

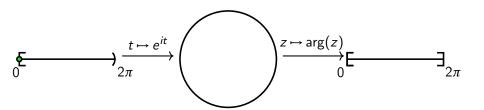
$$^{\dagger}f_n(\cdot,\delta_n)\circ\cdots\circ^{\dagger}f_1(\cdot,\delta_1)(x)\subseteq B(f_n\circ\cdots\circ f_1(x),\varepsilon)?$$

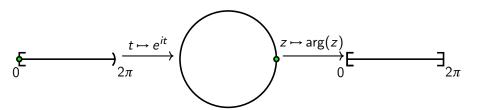
- **1** Obvious necessary condition: $f_n \circ \cdots \circ f_1$ continuous at x.
- Obviously not sufficient.



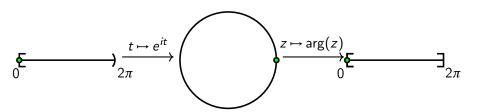




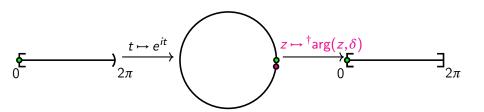




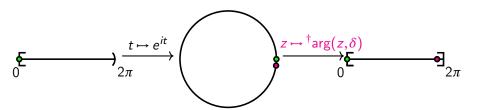
Example



Example



Example



The lattice of compacts

- Let Y be an effective metric space.
- ② Let $\mathcal{K}(Y)$ be the space of compact subsets of Y with the upper Vietoris topology.
- **3** Let $\mathcal{K}_{\perp}(Y)$ be the same space with a bottom element added.
- Then $\mathcal{K}_{\perp}(Y)$ is a complete lattice. Effectively so:

$$\bigcup \colon \mathscr{K}(\mathscr{K}_{\perp}(Y)) \to \mathscr{K}_{\perp}(Y)$$

and

$$\bigcap : \mathscr{V}(\mathscr{K}_{\perp}(Y)) \to \mathscr{K}_{\perp}(Y)$$

are continuous.

The lattice of compacts

Every function

$$f: X \to \mathscr{K}_{\perp}(Y)$$

has a best continuous approximation

$$F: X \to \mathscr{K}_{\perp}(Y),$$

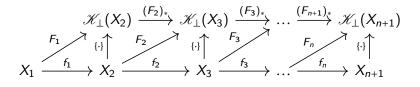
i.e.,

- ② If $G: X \to \mathcal{K}_{\perp}(Y)$ is continuous with $f(x) \in G(x)$ for all $x \in X$ then $F(x) \subseteq G(x)$ for all $x \in X$.

 \mathcal{K}_{\perp} is a monad, yielding a natural notion of composition for functions $F: X \to \mathcal{K}_{\perp}(Y), \ G: Y \to \mathcal{K}_{\perp}(Z)$. Explicitly:

$$G \circ F(x) = \bigcup_{y \in F(x)} G(y).$$

Overview



Theorem

Let $F_i: X_i \to \mathscr{K}_{\perp}(X_{i+1})$ be the best continuous approximation of $f_i: X_i \to X_{i+1}$. Assume that $F_i(x) \neq \bot$ for all $x \in X_i$. Then the following are equivalent:

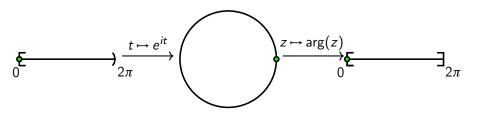
0

$$\forall \varepsilon > 0. \forall x \in X_1. \exists \delta > 0.$$

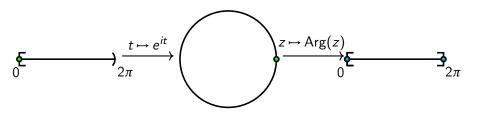
$$\left({}^{\dagger} f_n(\cdot, \delta) \circ \cdots \circ {}^{\dagger} f_1(\cdot, \delta)(x) \subseteq B(f_n \circ \cdots \circ f_1(x), \varepsilon). \right).$$

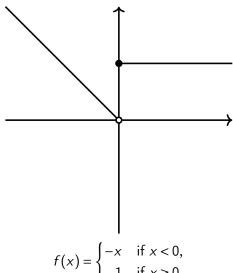
- **2** There is a continuous witness $\omega: X_1 \times \mathbb{Q}_{>0} \leadsto \mathbb{Q}_{>0}$ for the above.

Confirming what we already know

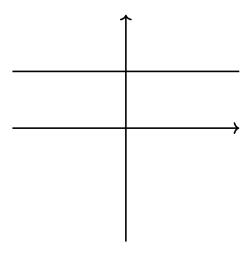


Confirming what we already know

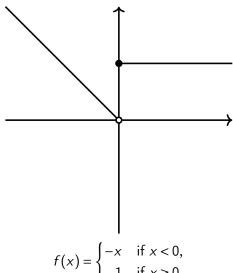




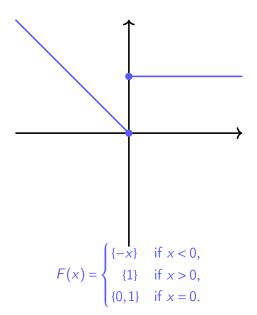
$$f(x) = \begin{cases} -x & \text{if } x < 0, \\ 1 & \text{if } x \ge 0. \end{cases}$$

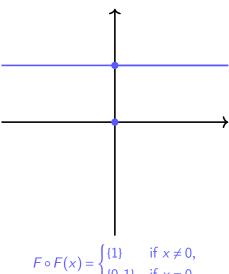


 $f\circ f(x)=1.$

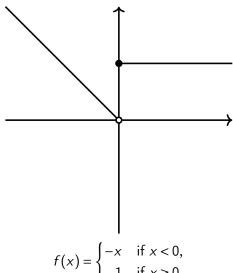


$$f(x) = \begin{cases} -x & \text{if } x < 0, \\ 1 & \text{if } x \ge 0. \end{cases}$$

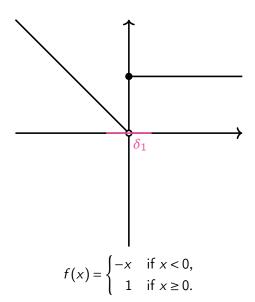


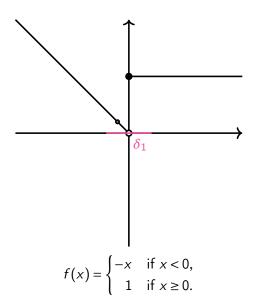


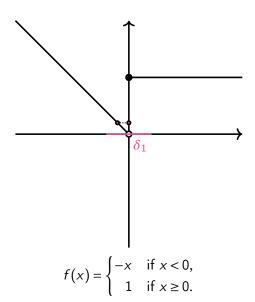
$$F \circ F(x) = \begin{cases} \{1\} & \text{if } x \neq 0, \\ \{0, 1\} & \text{if } x = 0. \end{cases}$$

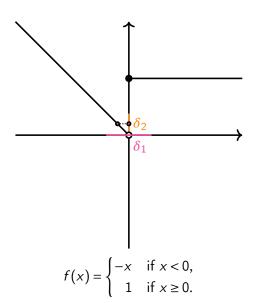


$$f(x) = \begin{cases} -x & \text{if } x < 0, \\ 1 & \text{if } x \ge 0. \end{cases}$$









Possible notions of convergence

1 A-priori uniform δ :

$$\forall \varepsilon > 0. \forall x \in X_1. \exists \delta > 0.$$

$$\left({}^{\dagger} f_n(\cdot, \delta) \circ \cdots \circ {}^{\dagger} f_1(\cdot, \delta)(x) \subseteq B(f_n \circ \cdots \circ f_1(x), \varepsilon). \right).$$

Adaptive scheme:

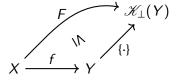
$$\forall \varepsilon > 0.$$

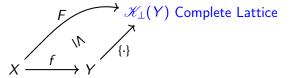
$$\forall x_1 \in X_1. \exists \delta_1 > 0.$$

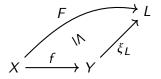
$$\forall x_2 \in f_1(x_1, \delta_1). \exists \delta_2 > 0.$$
....
$$\forall x_n \in f_{n-1}(x_{n-1}, \delta_{n-1}). \exists \delta_n > 0.$$

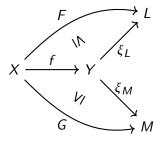
$$\forall x_{n+1} \in f_n(x_n, \delta_n).$$

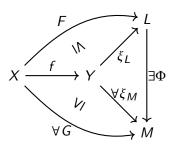
$$(x_{n+1} \in B(f_n \circ \cdots \circ f_1(x_1), \varepsilon)).$$



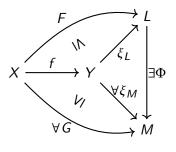








- 1. $\Phi \circ \xi_L = \xi_M$
- 2. $\Phi \circ F \geq G$



- 1. $\Phi \circ \xi_L = \xi_M$
- 2. $\Phi \circ F \geq G$

 \leadsto Existence of Φ requires injectivity.

Short comment on injective spaces

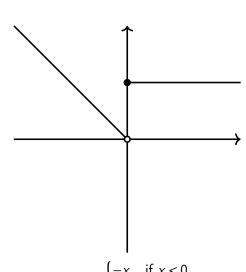
- Injective topological spaces w.r.t. subspace embeddings = continuous lattices.
- 2 Subspaces in QCB₀ carry sequentialisation of relative topology.
- \odot \Rightarrow Σ is not injective w.r.t. subspace embeddings.
- **3** Solution: consider injectivity with respect to Σ -split embeddings.
- **⑤** $i: X \to Y$ is Σ-split if $i^*: \mathcal{O}(Y) \to \mathcal{O}(X)$ admits a continuous section $s: \mathcal{O}(X) \to \mathcal{O}(Y)$.
- **⊙** Injective spaces are closed under retracts, finite products, and form an exponential ideal in the category of represented spaces. \Rightarrow For all represented spaces X the space $\mathcal{O}^2(X)$ is injective.
- o Injective spaces are simultaneously \mathcal{K} -algebras and \mathcal{V} -algebras, and in particular complete lattices.

- **1** Any function $f: X \to Y$ between represented spaces has a universal envelope.
- ② There exists an injective space \mathfrak{A}_f (unique up to unique isomorphism) and a continuous surjection

$$\pi:\mathfrak{A}_f\to\mathscr{O}(Y)$$

such that π preserves arbitrary joins and the fibres $\pi^{-1}(U)$ are injective spaces for all $U \in \mathcal{O}(Y)$.

- **③** The fibres $\pi^{-1}(U)$ encode extra information that is needed to verify for a given $x \in X$ if $f(x) \in U$.
- **1** The universal envelope is given by the best continuous approximation of f of type $X \to \mathcal{O}(\mathfrak{A}_f)$.
- § Simplest case $\mathfrak{A}_f = \mathscr{O}(Y)$ and $\pi = \mathrm{id}_{\mathscr{O}(Y)}$. For Hausdorff Y we recover best continuous approximations in $\mathscr{K}_{\perp}(Y)$.



Fibre of $U \in \mathcal{O}(\mathbb{R})$ non-trivial \Leftrightarrow $1 \in U$, $0 \notin U$, $(0,\delta) \subseteq U$ for some δ .

 $\pi^{-1}(U) = [0,\delta)_{<}/\sim,$ where $\delta = \sup_{(0,\eta) \subseteq U} \eta$, and \sim identifies all numbers > 0.

$$F(x)(U,[\delta]) = T$$

$$\Leftrightarrow$$

$$(x \neq 0 \land f(x) \in U)$$

$$\lor (1 \in U \land \exists \eta > 0.B(0,\eta) \subseteq U \land |x| < \eta)$$

$$\lor (1 \in U \land |x| < \delta)$$

$$F\colon X\to \mathcal{O}\left(\sum_{U\in\mathcal{O}(Y)}\mathcal{O}(\pi_f^{-1}(U))\right),\ G\colon Y\to \mathcal{O}\left(\sum_{V\in\mathcal{O}(Z)}\mathcal{O}(\pi_g^{-1}(V))\right)$$

$$F \circ G(x \in X) \left(V \in \mathcal{O}(Z), \alpha \in \pi_g^{-1}(V) \right)$$

= ???

$$F\colon X\to \mathcal{O}\left(\sum_{U\in\mathcal{O}(Y)}\mathcal{O}(\pi_f^{-1}(U))\right),\ G\colon Y\to \mathcal{O}\left(\sum_{V\in\mathcal{O}(Z)}\mathcal{O}(\pi_g^{-1}(V))\right)$$

$$F \circ G(x \in X) \left(V \in \mathcal{O}(Z), \alpha \in \pi_g^{-1}(V) \right)$$

= $F(x \in X)$ (????,???)

$$F\colon X\to \mathcal{O}\left(\sum_{U\in\mathcal{O}(Y)}\mathcal{O}(\pi_f^{-1}(U))\right),\ G\colon Y\to \mathcal{O}\left(\sum_{V\in\mathcal{O}(Z)}\mathcal{O}(\pi_g^{-1}(V))\right)$$

$$F \circ G(x \in X) \left(V \in \mathcal{O}(Z), \alpha \in \pi_g^{-1}(V) \right)$$

= $F(x \in X) \left(U = \left\{ y \in Y \mid G(y, V, \alpha) = \top \right\} \in \mathcal{O}(Y), ???? \right)$

$$F\colon X\to \mathcal{O}\left(\sum_{U\in\mathcal{O}(Y)}\mathcal{O}(\pi_f^{-1}(U))\right),\ G\colon Y\to \mathcal{O}\left(\sum_{V\in\mathcal{O}(Z)}\mathcal{O}(\pi_g^{-1}(V))\right)$$

$$F \circ G(x \in X) \left(V \in \mathcal{O}(Z), \alpha \in \pi_g^{-1}(V) \right)$$

= $F(x \in X) \left(U = \left\{ y \in Y \mid G(y, V, \alpha) = \top \right\} \in \mathcal{O}(Y), \top \in \pi_f^{-1}(U) \right)$

$$F\colon X\to \mathcal{O}\left(\sum_{U\in\mathcal{O}(Y)}\mathcal{O}(\pi_f^{-1}(U))\right),\ G\colon Y\to \mathcal{O}\left(\sum_{V\in\mathcal{O}(Z)}\mathcal{O}(\pi_g^{-1}(V))\right)$$

$$F \circ G(x \in X) \left(V \in \mathcal{O}(Z), \alpha \in \pi_g^{-1}(V) \right)$$

$$= F(x \in X) \left(U = \left\{ y \in Y \mid G(y, V, \alpha) = T \right\} \in \mathcal{O}(Y), T \in \pi_f^{-1}(U) \right)$$
No longer continuous in the open set argument.

$$F\colon X\to \mathcal{O}\left(\sum_{U\in\mathcal{O}(Y)}\mathcal{O}(\pi_f^{-1}(U))\right),\ G\colon Y\to \mathcal{O}\left(\sum_{V\in\mathcal{O}(Z)}\mathcal{O}(\pi_g^{-1}(V))\right)$$

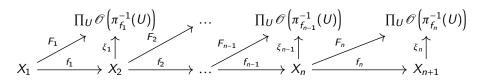
$$F \circ G(x \in X) \left(V \in \mathcal{O}(Z), \alpha \in \pi_g^{-1}(V) \right)$$

$$= F(x \in X) \left(U = \left\{ y \in Y \mid G(y, V, \alpha) = T \right\} \in \mathcal{O}(Y), T \in \pi_f^{-1}(U) \right)$$
No longer continuous in the open set argument.

This yields a function

$$G \circ F : X \to \prod_{V \in \mathscr{O}(Z)} \mathscr{O}\left(\pi_g^{-1}(V)\right)$$

Overview



Theorem

The following are equivalent:

0

$$\forall \varepsilon > 0.$$

$$\forall x_1 \in X_1.\exists \delta_1 > 0.$$

$$\forall x_2 \in {}^{\dagger}f_1(x_1, \delta_1).\exists \delta_2 > 0.$$
....
$$\forall x_n \in {}^{\dagger}f_{n-1}(x_{n-1}, \delta_{n-1}).\exists \delta_n > 0.$$

$$\forall x_{n+1} \in {}^{\dagger}f_n(x_n, \delta_n).$$

$$(x_{n+1} \in B(f_n \circ \cdots \circ f_1(x_1), \varepsilon)).$$

- **2** We can find continuous witnesses $\omega_i: X_i \times \mathbb{Q}_{>0} \rightsquigarrow X_{i+1}$ for the above.

Conclusion/Future Work

- Connected "upper" and "lower" relaxations of "exact" computational problems.
- **2** Replaces "hard analysis" questions on ε 's and δ 's by "soft analysis" questions on equality.
- **3** Envelopes can serve as a foothold for proving quantitative ε - δ -results.
- Backward approximations can yield efficient implementations of programs using envelopes.

TODOs:

- Extend results to WHILE-programs (mostly done).
- Extend to non-deterministic functions (partially done).
- **3** Pursue extraction of bounds for δ 's from equality proofs (promising preliminary results).
- Pipe dream: write programs with envelope semantics, prove them correct, compile program + proof into efficient program using backward approximations (entirely delusional).