Layout Randomization and Nondeterminism

Martín Abadi Jérémy Planul Gordon Plotkin

PrakashFest, Oxford, May 2014

Hoping to work with Prakash: Choquet capacities

Josée, Vineet, Radha, and Prakash, CONCUR '02: Weak Bisimulation is Sound and Complete for pCTL* .

• Superadditive Choquet capacity:

 $\mu(A \cup B) \le \mu(A) + \mu(B)$ (A, B $\subseteq \Sigma$, disjoint)

• Given measures μ_i , their lower envelope:

$$\mu(A) =_{def} \min_{i} \mu_i(A)$$

is a superadditive Choquet capacity.

• Can we do ordinary + probabilistic nondeterminism this way?

Hoping to work with Prakash: Choquet capacities

Josée, Vineet, Radha, and Prakash, CONCUR '02: Weak Bisimulation is Sound and Complete for pCTL* .

• Superadditive Choquet capacity:

 $\mu(A \cup B) \le \mu(A) + \mu(B)$ (A, B $\subseteq \Sigma$, disjoint)

• Given measures μ_i , their lower envelope:

$$\mu(A) =_{def} \min_{i} \mu_i(A)$$

is a superadditive Choquet capacity.

- Can we do ordinary + probabilistic nondeterminism this way?
- Sadly, this map is many-to-one from convex sets of measures to capacities, even when |Σ| = 3.

Barycentric coordinates



Counterexamples in plenty (Cozman)



Trying to work with Prakash: quantum computation

A Language and Type Theory for Quantum Computing

Prakash Panangaden* School of Computer Science McGill University Montreal, Quebec, Canada Gordon D. Plotkin[†] Department of Computer Science University of Edinburgh Edinburgh, Scotland, U. K.

February 28, 2003

Abstract

The abstract is very uncertain.

1 Introduction

Outline the nature of quantum comp as a web of interactions Our philosophy: evolution is fundamental, density ops, intervention operators are fundamental, no relativistic effects, entanglement, expressing unitaries; cite Peres and NC

Computation is dynamics. One prepares a dynamical system in some initial state - implicitly or explicitly containing the "imput" - and lets it evolve. The "answer" is contained in the final state. Understanding quantum computation must, at its lowest level, involve understanding quantum evolution.

Normally in quantum mechanics understanding quantum evolution means understanding the generator of time evolution, i.e. the Hamiltonian operator. If H is the Hamiltonian operator for a system than the evolution of states is governed by the *unitary* operator e^{-HH} . Typically this is cast as a differential equation, the Schnedinger equation. In typical text books one spends a significant amount of time solving the Schnedinger equation for a writely of Hamiltonians.

In quantum computation textbooks [NO0] one takes varians unitary operations on states of collections two-state systems (unlike) as given. These tox-state systems: or quibits, are concived of as separate carriers of information, just like classical bits, Physically they are generally realized as spin-states of spin $\frac{1}{2}$ particles or as polarization states of photons. What is very different about quibits versus bits is that (1) a quibit may exist in a superspect state i.e. in a linear combination of its two basic states and (ii) states of different quibits may be entangled i.e. correlated in nontrivial ways.

Algorithms are developed by combining the primitive unitary operators and occasionally by making measurements. However, the primitive unitaries do not all just act on one qubit at a time.

Somewhat working with Prakash: Markov processes

Approximating Markov Processes By Averaging

PHILIPPE CHAPUT McGill University VINCENT DANOS University of Edinburgh PRAKASH PANANGADEN McGill University GORDON PLOTKIN University of Edinburgh

October 14, 2013

Abstract

Normally, one thinks of probabilistic transition systems as taking an initial probability distribution over the state space into a new probability distribution representing the system after a transition. We, however, take a dual view of Markov processes as transformers of bounded measurable functions. This is very much in the same spirit as a "predicate-transformer" view, which is dual to the state-transformer view of transition systems.

We redevelop the theory of labelled Markov processes from this view point, in particular we explore approximation theory. We obtain three main results:

(i) It is possible to define bisimulation on general measure spaces and show that it is an equivalence relation. The logical characterization of bisimulation can be done straightforwardly and generally. (ii) A new and flexible approach to approximation based on averaging can be given. This vastly generalizes and streamlines the idea of using conditional expectations to compute approximations. (iii) We show that there is a minimal process bisimulation-equivalent to a given process, and this minimal process is obtained as the limit of the finite approximatis.

Low-Level Attacks and Protection

- Many attack techniques:
 - Buffer overflows
 - Exception overwrites
 - return-to-libc
 - jump-to-libc
 - use-after-free attacks
- Often with knowledge and control of the heap

- Many defenses:
 - Stack canaries
 - Safe exception handling
 - NX (No eXecute) data
 - Layout randomization
- Useful mitigations
- But not necessarily perfect in a precise sense
- Nor all well understood

Runtime attacks often depend on addresses (e.g. jump to libc).

- Let us randomize the addresses!
 - Considered for data at least since the rise of large virtual address spaces (e.g., [Druschel & Peterson, 1992] on fbufs).
 - Now present in Linux (PaX), Windows, Mac OS X, iOS, Android (4.0).



A theory of layout randomization

- Abadi and Plotkin did a first formalization, using operational methods.
- Further work [Jagadeesan et al., Abadi et al.] concerned memories containing functions and arrays.
- Here we consider commands that make nondeterministic choices (as a first step towards parallelism).
- We use denotational methods rather than operational ones (for interest, but not only).
- We use random variables for probability rather than distributions, as events are not independent.

Probabilities and nondeterminism

Probabilities:



Not red with ≥ 0.9 probability

Never red

- Probabilistic cryptographic schemes
- Differential privacy
- Low-level software protection (Layout randomization)

- Network communications
- Nondeterministic scheduling
- User interaction

Probabilities ↔ Nondeterminism: Game theory

- Player 1 chooses a bit (b₁) by flipping a coin
- Player 2 chooses a bit (b₂) nondeterministically
- Player 1 wins if he guesses right:

$$b_1 = b_2$$

• The program:

$$b_1 := 0 + \frac{1}{2} 1; (b_2 := 0 \text{ OR } b_2 := 1)$$

Probabilities ↔ Nondeterminism: Game theory

$$b_1 := 0 + \frac{1}{2} 1; (b_2 := 0 \text{ OR } b_2 := 1)$$

What should its semantics be?

- a set of distributions on ٠ outcomes 0.5 0.5 0.5 0.5
- Player I always wins with probability 1/2.

outcomes

a distribution on sets of

Player I can win or lose.

$$b_1 := 0 + \frac{1}{2} 1; (b_2 := 0 \text{ OR } b_2 := 1)$$

The obvious semantics for sets of distributions on outcomes does not compose as expected:



Probabilities ↔ Nondeterminism: Algebra

- Resolving probabilistic choice first:
 - Distribute nondeterministic choice over probabilistic choice:

$$x \cup (y +_{\rho} z) = (x \cup y) +_{\rho} (x \cup z)$$

- One can prove that for all p, q: x +_p y = x +_q y so the +_p become a second semilattice, distributing over the first one.
- Free Algebra over a set *X*: Non-empty finite ∪-closed sets of non-empty finite subsets of *X*.
- Resolving nondeterministic choice first:
 - Distribute probabilistic choice over nondeterministic choice:

$$x+_{\rho}(y\cup z)=(x+_{\rho}y)\cup(x+_{\rho}z)$$

- Free Algebra over a set *X*: Non-empty finitely generated convex sets of distributions over *X*.
- But gives the "wrong" semantics to our program.

Roadmap

- Define high-level and low-level languages, each with access to (the same) private and public locations.
 - \Rightarrow View the implementations as compilations.
- View attackers as (high- or low-level) "public" contexts: those with no access to private locations.

 \Rightarrow Relate commands depending on their public behaviour in public contexts.

 Characterize the contextual relations using semantical simulations (a kind of logical relation)

 \Rightarrow prove that high-level contextual relations correspond to low-level contextual relations (full abstraction):

$$\begin{array}{ccc} \llbracket c \rrbracket \preceq \llbracket c' \rrbracket & \Leftrightarrow & \forall \text{high-level public } C[\]. \llbracket C[c] \rrbracket \leq_L \llbracket C[c'] \rrbracket \\ & \updownarrow \\ \llbracket c^{\downarrow} \rrbracket \preceq \llbracket c'^{\downarrow} \rrbracket & \Leftrightarrow & \forall \text{low-level public } C[\]. \llbracket C[c^{\downarrow}] \rrbracket \leq_L \llbracket C[c'^{\downarrow}] \rrbracket \end{array}$$

High-level language

$$e ::= k | !I_{loc} | e + e | e * e$$

$$b ::= e \le e | \neg b | true | false | b \lor b | b \land b$$

$$c ::= I_{loc} := e | if b then celse c | skip | c; c |$$

$$c + c | while b do c$$

- Finitely many locations.
- They are divided into private (high) and public (low).
- store s: locations → values
- Semantics [[c]]: stores $\rightarrow \mathscr{H}(\text{stores}_{\perp})$

$$e ::= k | !I_{loc} | e + e | e * e$$

$$b ::= e \le e | \neg b | true | false | b \lor b | b \land b$$

$$c ::= I_{loc} := e | if b then celse c | skip | c; c |$$

$$c + c | while b do c$$

- Finitely many locations.
- They are divided into private (high) and public (low).
- store s: locations → values
- Semantics [[c]]: stores $\rightarrow \mathscr{H}(\text{stores}_{\perp})$

Refinement relation (high-level)

For X, Y ∈ ℋ(stores_⊥), X refines Y if the public part of any store in X is the public part of a store in Y.



Formally:

$$X \leq_L Y \equiv \forall s \in \text{stores.} s_L \in X \Rightarrow s_L \in Y$$

For commands, *c* refines *c*′, written *c* ≤_{*L*} *c*′ when, for every input, the outputs of *c* refine those of *c*′:

$$\forall s \in \text{stores.} \llbracket c \rrbracket(s) \leq_L \llbracket c' \rrbracket(s)$$

Simulation (high-level)

• Simulation relation \preceq on commands such that

Theorem 1

For all high-level commands c, c', we have:

 $\llbracket c \rrbracket \preceq \llbracket c' \rrbracket \iff \forall \text{ high-level public } C \llbracket . \llbracket C [c] \rrbracket \leq_L \llbracket C [c'] \rrbracket$

• Simple without nondeterminism or nontermination:



Simulation (details)



- So begin with $R \subseteq \mathcal{H}(\text{private-stores}_{\perp})^2$, subject to conditions
- Then construct $R^+ \subseteq \mathscr{H}(\text{stores}_{\perp})^2$ from R
- Then

 $\llbracket c \rrbracket \preceq \llbracket c' \rrbracket \equiv_{def} \forall X, Y \in \mathscr{H}(\text{stores}_{\perp}). XR^+ Y \Rightarrow \llbracket c \rrbracket (X)R^+ \llbracket c' \rrbracket (Y)$

Low-level memory model

- Memories (*m*∈ Mem) are partial mappings from the set of memory addresses Add = {1,...,*r*} to values.
- Memory layouts (w ∈ Lay) are injective mappings from locations to the addresses Add.
- \implies stores map to memories via layouts.



Layout randomization

- We fix a public layout w_p.
- We randomize the private layout (using a distribution *d*)
- We rely on one value in our theorems, the minimum probability for a guessed private address to be unallocated:

$$\delta = \min\{P(i \notin \operatorname{ran}(w)) \mid i \in \operatorname{Add} \setminus \operatorname{ran}(w_p)\} \\ = 1 - |\operatorname{PriLoc}|/|\operatorname{Add} \setminus \operatorname{ran}(w_p)| \quad (d \text{ uniform})$$



$$e ::= k |I_{nat}|!e|e+e|e*e$$

$$b ::= e \le e |\neg b| true| false|b \lor b|b \land b$$

$$c ::= e := e | if b then celse c | skip | c; c |$$

$$c + c | while b do c$$

• We can compute over I_{nat} , ex: $(x + I_{\text{nat}}) :=_{\text{nat}} (x + I'_{\text{nat}})$.

$$e ::= k |I_{nat}|!e|e+e|e*e$$

$$b ::= e \le e |\neg b| true| false|b \lor b|b \land b$$

$$c ::= e := e | if b then celse c | skip | c; c |$$

$$c + c | while b do c$$

• We can compute over I_{nat} , ex: $(x + I_{\text{nat}}) :=_{\text{nat}} (x + I'_{\text{nat}})$.

Low-level semantics

- The final memory depends on the initial one and the layout,
- We cannot prove security using a semantics of the form:

$$\operatorname{Mem} \times \operatorname{Lay} \to \mathscr{H}(\operatorname{Mem}_{\xi,\perp})$$

for consider the command:



Martín Abadi, Jérémy Planul, Gordon Plotkir

- So we need to resolve probabilities after nondeterminism:
- However

$$\operatorname{Mem} \to \mathscr{H}(\mathscr{V}(\operatorname{Mem}_{\xi,\perp}))$$

does not work as our events are not independent.

- For example, suppose there is just one private location *I*₁, and 1,2 are not public addresses. Then:
 - 1:=1; 2:=1 should always give an error.
 - 1 := 1; 1 := 2 should give an error or overwrite l_1

• Solution idea: replace distributions

 $d \in \mathscr{V}(\operatorname{Mem}_{\xi,\perp})$

by random variables

 ζ : Lay \rightarrow Mem_{ξ,\perp}

with layouts as the sample space, obtaining:

$$\operatorname{Mem} \to \mathscr{H}(\operatorname{Lay} \to \operatorname{Mem}_{\xi, \perp})$$

which compose, using Kleisli structure on \mathcal{H} .

Low-level semantics

Solution idea: replace distributions

 $d \in \mathscr{V}(\operatorname{Mem}_{\xi,\perp})$

by random variables

$$\zeta$$
: Lay \rightarrow Mem _{ξ,\perp}

with layouts as the sample space, obtaining:

$$\operatorname{Mem} \to \mathscr{H}(\operatorname{Lay} \to \operatorname{Mem}_{\xi, \perp})$$

• Sadly, these do not compose (cf intro) and we instead use:

$$(\mathsf{Lay} \to \mathrm{Mem}_{\xi,\perp}) \to \mathscr{H}(\mathsf{Lay} \to \mathrm{Mem}_{\xi,\perp})$$

which compose, using Kleisli structure on \mathcal{H} .

Low level semantics

• Example:



So, as desired, always get an error with high probability.

$$\llbracket c \rrbracket : (\mathsf{Lay} \to \mathrm{Mem}_{\xi, \bot}) \to \mathscr{H}(\mathsf{Lay} \to \mathrm{Mem}_{\xi, \bot})$$

Most cases are straightforward:

$$\begin{split} & \llbracket c + c' \rrbracket(\zeta) &= \llbracket c \rrbracket(\zeta) \cup \llbracket c' \rrbracket(\zeta) \\ & \llbracket c; c' \rrbracket &= \llbracket c' \rrbracket^{+} \circ \llbracket c \rrbracket \\ & \llbracket skip \rrbracket &= \eta \\ & \llbracket e := e' \rrbracket(\zeta) &= \begin{cases} \eta(\lambda w. m[\llbracket e \rrbracket_{\zeta(w)}^w \mapsto \llbracket e' \rrbracket_{\zeta(w)}^w]) \\ & (\text{if } \llbracket e \rrbracket_{\zeta(w)}^w \in \text{dom}(m)) \\ & (\text{if } \llbracket e \rrbracket_{\zeta(w)}^w \in \text{dom}(m)) \end{cases} \\ & \\ & [\texttt{if } b \texttt{then } c \texttt{else } c' \rrbracket = \texttt{Cond}(\llbracket b \rrbracket, \llbracket c \rrbracket), \llbracket c' \rrbracket) \\ & [\texttt{while } b \texttt{do } c \rrbracket &= \mu \theta. \texttt{Cond}(\llbracket b \rrbracket, \theta^{+} \circ \llbracket c \rrbracket, \eta) \end{split}$$

Example conditional semantics



Refinement relation (low-level)

- Input We are interested in the distinguishing capacities of the adversary on inputs that directly correspond to stores. These are called store projections and are random variables (Lay → Mem_{⊥,ξ}) of the form w → w(s) for some store s (but also allow ⊥ and ξ).
- Output For X, Y ∈ ℋ(Lay → Mem_{⊥,ζ}), X ≤_L Y iff (roughly) the public parts of the random variables in X are less than those of Y with probability greater than δ. More precisely, for all ζ ∈ X, there exists ζ' ∈ Y such that:
 - $\zeta_L \leq \zeta'_L$, or • $P(\zeta(w) \in \{\xi, \bot\}) \geq \delta$ and $P(\zeta'(w) = \xi) \geq \delta$.
- Command refinement *c* refines c', written $c \leq_L c'$, when:

$$\forall s \in \text{stores.} \llbracket c \rrbracket (w \mapsto w(s)) \leq_L \llbracket c' \rrbracket (w \mapsto w(s))$$

Simulation relation

- We mostly consider public contexts containing safe commands.
- Safe commands transform store projections into store projection sets, which are (downsets of) sets of
 - store projections, or
 - random variables with high probability of error (> δ).
- We project those semantics and define a simulation relation similar to the high-level one.

Theorem 2

For all safe low-level commands c, c', we have:

 $\llbracket c \rrbracket \preceq \llbracket c' \rrbracket \iff \forall \text{ low-level public } C[\]. \llbracket C[c] \rrbracket \leq_L \llbracket C[c'] \rrbracket$

Simulation relation - more detail

- Begin with $R \subseteq \mathscr{H}(\text{private-stores}_{\perp,\xi})^2$, subject to conditions
- Then construct $R^+ \subseteq \mathscr{H}(\text{stores}_{\perp,\xi})^2$ from R, much as before
- Then construct $R^{\times} \subseteq \mathscr{H}(\text{Lay} \to \text{Mem}_{\xi,\perp})^2$ from R^+ by:

$$XR^{\times}Y \equiv_{def} X \text{ and } Y \text{ are store projection sets}$$

 $\wedge \chi(X)R^{+}\chi(Y)$

(where χ extracts the stores, bottom, and, possibly, error from a store projection set).

Then

$$\llbracket c \rrbracket \preceq \llbracket c' \rrbracket \equiv_{def} \forall X, Y. XR^{\times} Y \Rightarrow \llbracket c \rrbracket (X)R^{\times} \llbracket c' \rrbracket (Y)$$

Compilation

High-level commands *c* are compiled to low-level commands c^{\downarrow} by:

$$(!I_{loc})^{\downarrow} = !I_{nat}$$

 $(I_{loc}:=e)^{\downarrow} = I_{nat}:=e^{\downarrow}$

Theorem 3

For all high-level commands c, c' we have: $c \preceq c'$ iff $c^{\downarrow} \preceq c'^{\downarrow}$

Theorem 4

For all high-level commands c, c',

$$\forall$$
 high-level public $C[].[[C[c]]] \leq_L [[C[c']]]$

holds iff

 \forall low-level public $C[].[[C[c^{\downarrow}]]] \leq_L [[C[c'^{\downarrow}]]]$

does

Naturality of the semantics

Parameterising over sets of layouts:



Suggests working with presheaves.

A semantic approach to layout randomization

$$\begin{split} \llbracket c \rrbracket \preceq \llbracket c' \rrbracket & \Leftrightarrow & \forall \text{ public } C[\]. \llbracket C[c] \rrbracket \leq_{L} \llbracket C[c'] \rrbracket \\ & \updownarrow \\ \llbracket c^{\downarrow} \rrbracket \preceq \llbracket c'^{\downarrow} \rrbracket & \Leftrightarrow & \forall \text{ public } C[\]. \llbracket C[c^{\downarrow}] \rrbracket \leq_{L} \llbracket C[c'^{\downarrow}] \rrbracket \\ \end{split}$$

- Nondeterminism (and its interaction with probabilities).
- Concurrency would be an interesting next step (or higher-order)

• A semantic approach to layout randomization

- Nondeterminism (and its interaction with probabilities).
- Concurrency would be an interesting next step (or higher-order)
- But there is still something to understand.

What is the corresponding operational semantics?

• To make sense of our semantics, we want operational/denotational consistency, akin to:

$$\zeta' \in \llbracket c \rrbracket(\zeta) \Leftrightarrow \forall w. w \models \langle c, \zeta(w) \rangle \to^* \zeta'(w)$$

- Once again, the nondeterminism is chosen after the layout.
- We set the nondeterminism with an oracle, then the operational semantics is deterministic:

$$\zeta' \in \llbracket c \rrbracket(\zeta) \iff \exists \pi. \forall w. w \models \langle c, \zeta(w), \pi \rangle \to^* \zeta'(w)$$