CS 7480 Spring 2013 Control Flow Analysis

notes by: Stephen Chang

Friday, January 11, 2013

1 Summary

General monotone flow analysis framework.

1. Monotonic, continuous functions
2. Lattices
   - represents a program property, ie, a "fact" lattice
   - ordering is information ordering
   - modeling programs as functions on lattices

2 Monotonic, continuous functions

Definition 1. *A* least upper bound (lub) *on set S:*

1. $\forall x \in S : x \leq \operatorname{lub} S$,
   ie, $\operatorname{lub} S$ *is an upper bound of* S
2. $\forall x \in S : x \leq y \implies \operatorname{lub} S \leq y$,
   ie, $\operatorname{lub} S$ *is less than or equal to all upper bounds of* S.

If ordering is information ordering, then $\operatorname{lub} S$ is the "most precise" upper bound.

Definition 2. *In a* complete partial order (cpo)*, every chain has a lub.*

Definition 3. *A* pointed set *has a bottom ($\perp$) element.*

Definition 4. *A function f is* monotone *if* $\forall a, b : a \leq b \implies f(a) \leq f(b)$.

In other words, more information gives better conclusions. But we don't need to worry much about monotonicity because it's pretty much always true.

Definition 5. *A function f is* continuous *if* $\forall$ *chains* $C : f(\operatorname{lub} C) = \operatorname{lub} f(C)$.

In other words, the function and lub operations commute. So you "can sneak up on an answer."

Theorem 1. *continuous* $\implies$ *monotone*

Proof. .

1. $x \leq y \implies \text{lub}\,\{x, y\} = y$
2. From the definition of continuous:

   $$f(\text{lub}\,\{x, y\}) = \text{lub}\, f(\{x, y\})$$

   Using step 1:

   $$f(y) = \text{lub}\, f(\{x, y\}) = \text{lub}\,\{f(x), f(y)\}$$

   so

   $$f(x) \leq f(y)$$

   meaning f is monotone.

□

Theorem 2. *monotonic* $\implies$ *continuous, when cpo chains are finite*

Proof. Consider (finite) chain $C = x_1 \leq \cdots \leq x_n$. Want to show:

$$f(\text{lub}\, C) = \text{lub}\, f(C)$$

1. $\text{lub}\, C = x_n$, so $f(\text{lub}\, C) = f(x_n)$
2. By monotonicity:

   $$f(x_1) \leq \cdots \leq f(x_n)$$

   so

   $$\text{lub}\, f(C) = f(x_n)$$

□

Theorem 3. *With an infinite chain, monotone* $\not\Longrightarrow$ *continuous.*

Proof. Counterexample:
$f : \mathbb{N}_\infty \to \mathbb{N}_\infty$, $f(i) = 0$, for $i \in \mathbb{N}$, $f(\infty) = 1$
Consider the chain $C = \mathbb{N}$. $f(\text{lub}\, C) = f(\infty) = 1$, but $\text{lub}\, f(C) = 0$. □

Theorem 4 (Least Fixed-point)**.** *If D is a pointed cpo and $f : D \to D$ is continuous, then f has a* least fixed-point (lfp)*,* $\text{fix}\, f = \text{lub}\,\{f^n(\bot) \mid n \geq 0\}$

Recall that a recursive function is a search for a fixed point. If there is a least fixed-point, then there is a "best answer".

Proof. Summary:

1. Show that $\{f^n(\bot) \mid n \geq 0\}$ has a lub.

2. Show that the lub is a fixed-point.
3. Show that the lub is a least fixed-point.

Proof:

1. (a) continuous $\implies$ monotone, so $\perp \leq f(\perp)$, since $\perp$ is $\leq$ everything

   (b) using the same reasoning, we get the chain:

   $$\perp \leq f(\perp) \leq f^2(\perp) \leq f^3(\perp) \leq \cdots$$

   Since $f^0(\perp) = \perp$, $\{f^n(\perp) \mid n \geq 0\}$ is a chain.

   (c) since D is complete, every chain has a lub so $\{f^n(\perp) \mid n \geq 0\}$ has a lub

2. (a) Let $\mathrm{lub}\,\{f^n(\perp) \mid n \geq 0\} = \mathrm{fix}\, f$. Then applying f to both sides:

   $$f(\mathrm{fix}\, f) = f(\mathrm{lub}\,\{f^n(\perp) \mid n \geq 0\})$$

   (b) By continuity:

   $$f(\mathrm{lub}\,\{f^n(\perp) \mid n \geq 0\}) = \mathrm{lub}\, f(\{f^n(\perp) \mid n \geq 0\}) = \mathrm{lub}\,\{f^n(\perp) \mid n \geq 1\}$$

   (c) $\{f^n(\perp) \mid n \geq 1\}$ is still a chain and removing the bottom element doesn't change the lub, so:

   $$\mathrm{lub}\,\{f^n(\perp) \mid n \geq 1\} = \mathrm{lub}\,\{f^n(\perp) \mid n \geq 0\} = \mathrm{fix}\, f$$

   So $f(\mathrm{fix}\, f) = \mathrm{fix}\, f = \mathrm{lub}\,\{f^n(\perp) \mid n \geq 0\}$ is a fixed-point.

3. (a) Suppose d' is a fixed-point so $d' = f(d')$. We know $\perp \leq d'$ and by monotonicity, $f(\perp) \leq f(d')$.

   (b) But $f(d') = d'$ so $f(\perp) \leq d'$.

   (c) Applying f again $f^2(\perp) \leq f(d') = d'$ so $f^2 \leq d'$. Since $f^n(\perp) \leq d'$, d' is an upper bound of $\{f^n(\perp) \mid n \geq 0\}$.

   (d) Since $\mathrm{fix}\, f$ is the least upper bound of $\{f^n(\perp) \mid n \geq 0\}$, $\mathrm{fix}\, f \leq d'$. Since $\mathrm{fix}\, f = \mathrm{lub}\,\{f^n(\perp) \mid n \geq 0\}$ is also a fixed-point, it must be the least fixed-point.

$\square$

Takeaway from least fixed-point theorem:

If I can set up a problem as recursive equations, ie, a search for a fixed-point, then if my fact space has certain properties (pointed, continuous (and thus monotone) complete partial order), then there is a "best" answer.

Fun fact (unrelated to this course): The fix function is also continuous.

3 Lattices

Definition 6. *A poset* $(S, \leq)$ *is a* lattice *iff* $\forall x, y \in S$*, there exists a unique meet* $x \wedge y$ *and join* $x \vee y$ *such that:*

1. $x \wedge y \leq x$
2. $x \wedge y \leq y$
3. $z \leq x \wedge z \leq y \implies z \leq x \wedge y$

In other words, the meet is the greatest lower bound, ie the infimum. Also:

1. $x \vee y \geq x$
2. $x \vee y \geq y$
3. $z \geq x \vee z \geq y \implies z \geq x \vee y$

In other words, the join is the least upper bound, ie the supremum.

Other kinds of lattices:

- A semi-lattice has either a supremum or infimum but not both.
- A bounded lattice has a $\perp$ and $\top$ element.
- In a complete lattice, every set has a meet and join, including infinite sets.
- So complete $\implies$ bounded and finite lattices are complete.

Definition 7 (Alternative (algebraic) lattice definition)**.** *A* lattice *is a set with meet and join operations such that they are:*

- *commutative*
- *associative*
- *absorption (ie* $x \vee (x \wedge y) = x$ *and* $x \wedge (x \vee y) = x$*)*

With these properties, we can prove the conditions from definition 6.

meet and join are idempotent:

- $x \vee x = x$
- $x \wedge x = x$
- $x \vee \perp = x$
- $x \wedge \top = x$

(Unrelated) homework: Use the word "palimpsest" in a conversation.

Bonus: Use the word "boustrophedon" in a conversation.

Non-guaranteed properties of lattices:

- distributivity: $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$
- well-foundedness: every chain is finite, ie may be able to compute it

Some concrete lattices:

- powerset of S, $\mathcal{P}(S)$ is a lattice, $\leq = \subseteq$ (subset), meet $= \cap$, join $= \cup$, $\perp = \{\}$, $\top = S$
- integers $\mathbb{Z}$ is an unbounded lattice, meet = min, join = max
- $\mathbb{Z}_{-\infty,\infty}$ is a bounded lattice (I forgot the analysis examples that use this lattice)
- $\mathbb{Z}_{\perp,\top}$ is a flat lattice
- logicians use the lattice of propositions: meet = and, join = or, $\leq = \Longrightarrow$
- range analysis lattice $\mathcal{P}(Z)$: represent sets of integers as ranges $[i, j]$

  examples:

  $[0, 7] \vee [5, 10] = [0, 10]$

  $[0, 5] \vee [10, 15] = [0, 15]$ (some precision was lost)

In flow analysis, there are 2 sources of "crap" (imprecision)

1. join operation introduces approximation
2. (for some reason I didnt write down a #2)