1 Optimization (Continued)

1.1 Flow Analysis

1.1.1 Another Flow Analysis Walkthrough

Consider the following program:

```
(fun (same_at vec1 vec2 i)
    (= (index vec1 i) (index vec2 i)))
```

We'll perform another flow analysis[1], again by starting from the beginning and going to the end. The information we'll keep track of are the *potential* tags. Our tags are now slightly more refined; in particular, we now have the set of tags,

$$\text{tag} := Z|\text{Pos}|\text{Neg}|B|V|\text{Nil}$$

Here, `Z` means the number zero, `Pos` means positive number, `Neg` means negative number. We also have `B` for boolean, `V` for vector, and `Nil` for nil. Once again, we let A represent the set of all possible tags. We also introduce $N = \{Z, \text{Pos}, \text{Neg}\}$ for numbers[2].

IR	`vec1`	`vec2`	`i`	t_0	t_1	`rax`
`start0: check isnonnilvec(vec1)`	$\to V$					
`start1: check isnum(i)`	V		$\to N$			
`start2: check i >= 0`	V		$N \to \{Z, \text{Pos}\}$			
`start3: check i < len(vec1)`	V		$\{Z, \text{Pos}\} \to \{Z, \text{Pos}\}$			
`start4: %t_0 <- vec1[i]`	V		$\{Z, \text{Pos}\}$	$\to A$		
`start5: check isnonnilvec(vec2)`	V	$\to V$	$\{Z, \text{Pos}\}$	A		
`start6: check isnum(i)`	V	V	$\{Z, \text{Pos}\}$	A		
`start7: check i >= 0`	.	.	.	.		
`start8: check i < len(vec2)`	.	.	.	.		
`start9: %t_1 <- vec2[i]`	.	.	.	.		
`start10:check sametag(%t_0, %t_1)`						
`start10:rax <- %t_0 == %t_1`						
`return rax`						

Note that, at `start7`, at our second pass, we can either reduce this to just checking `true`, or just deleting the check altogether.

1.1.2 Abstract Domains

We're now incorporating forward analysis with some numeric range analysis. We can call these types of analysis abstract domains. We have seen three different abstract domains for analysis:

- `Set<String>` for liveliness analysis (for register allocation).
- `Dict<String, Set<Tag>>` for forward data analysis.

There were two different types of tags:

- `Z | P | N`: information about numbers (positive, negative, etc.)
- `N | B | Nil | P`: other relevant tag information.

How do we unify these domains? We can consider ideas like:

- Frequency for variable use.
- Frequency of branches.
- Ranges of numbers.
- Booleans as true or false.
- Lengths of vectors (if we have constants, e.g., setting the tag to be the length of the vector).

[1] Note that this IR representation was generated by hand.
[2] Note that there's no overlaps in this set; we either have negative numbers, positive numbers, and zero.