Notes on navigation history

Connor G. Brewster
Alan Jeffrey

DRAFT of 2016-06-24

Abstact: Navigation has been a core component of the web since its inception: users and scripts can follow hyperlinks, or can go back or forwards through the session history. Despite its long existence, navigation history has no formal model or properties. In this paper, we present a formal model aligned with the WHATWG specification of navigation history, and investigate its properties. The fundamental property of navigation is *homomorphism*: traversing the history by δ then by δ' should be the same as traversing by $\delta+\delta'$. In particular, traversing by $+1$ (forward) then by -1 (back) is the same as traversing by 0 (doing nothing). We show that the specification-aligned model does not satisfy this property, by exhibiting a series of counter-examples, which motivate three patches to the model. The patched model is homomorphic. We present a series of experiments, showing that browsers are inconsistent in their implementation of navigation history, but that their behaviour is closer to the patched model than to the specification-aligned model. We propose patches to the specification to align it with the homomorphic model.

ACM Classification: D.2.1 Requirements/Specifications.

Keywords: Formal model, Navigation, Session history, Specification, Web browsers.

1. Introduction

[These are rough notes, working towards a model of navigation history for the web.]

2. Preliminaries

[Define forest, tree, root, total order, equivalence.]

3. Model

A *navigation history* $H = (D, A, \rightarrow, \leq, \sim)$ consists of:

- a set D (the *documents*),
- a subset $A \subseteq D$ (the *active* documents),
- a forest $(D, \rightarrow)$ (the *document hierarchy*),
- a total order $(D, \leq)$ (the *chronological order*), and
- an equivalence relation $(D, \sim)$ (the *same-session equivalence*).

such that:

- for every d there is a unique $d' \in A$ such that $d \sim d'$,
- for every $d \rightarrow e \sim e'$ we have $d \rightarrow e'$, and

- for every $d \to e$, we have $d \leq e$.

Define:

- d_0 is the unique active root document,
- $d \twoheadrightarrow e$ when $d \to e$ and $e \in A$,
- $FA = \{d \mid d_0 \twoheadrightarrow^* d\}$ (the *fully active* documents),
- $d \lesssim e$ whenever $d \sim e$ and $d < e$,
- the *session future* of d is $\{e \mid d \lesssim e\}$,
- the *session past* of d is $\{e \mid d \gtrsim e\}$,
- the *joint session future* is $\{e \mid \exists d \in FA \,.\, d \lesssim e\}$,
- the *joint session past* is $\{e \mid \exists d \in FA \,.\, d \gtrsim e\}$,

Define *deleting d from H*, when $d \notin FA$, to be H' where:

- $D' = d \setminus \{e \mid d \to^* e\}$,
- $e \in A'$ whenever $e \in A$,
- $e \leq' f$ whenever $e \leq f$,
- $e \to' f$ whenever $e \to f$, and
- $e \sim' f$ whenever $e \sim f$.

Define *replacing d by d' in H*, where $d \in A$ and $d' \notin D$, to be H' where:

- $D' = D \cup \{d'\}$,
- $e \in A'$ whenever $e \in A$ and $e \neq d$, or $e = d'$,
- $e \leq' f$ whenever $e \leq f$, or $f = d'$,
- $e \to' f$ whenever $e \to f$, or $e \to d$ and $f = d'$, and
- $e \sim' f$ whenever $e \sim f$, or $e \sim d$ and $f = d'$, or $d \sim f$ and $e = d'$.

Define *navigating from d to d' in H* to be the result of:

- deleting the session future of d, and
- replacing d by d'.

Define *traversing the history to d in H* to be H' where:

- D' is D,
- $e \in A'$ whenever $d \not\sim e \in A$, or $d = e$,
- $e \leq' f$ whenever $e \leq f$,
- $e \to' f$ whenever $e \to f$, and
- $e \sim' f$ whenever $e \sim f$.

Define *H traverses the history by $+\delta$ to H'* when:

- the joint session future of H is $d_1 > \cdots > d_\delta > \cdots$,
- H traverses the history to d_δ in H'

Define H *traverses the history by* $-\delta$ *to* H' when:

- the joint session past of H is $d_1 < \cdots < d_\delta < \cdots$,
- H traverses the history to d_δ in H'

Define H *traverses the history by* 0 *to* H' when $H = H'$.

[This defn is meant to align with the spec.]

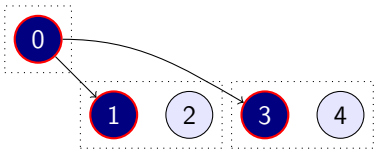
4. Properties

[State some goals, e.g. go(δ);go(δ') is the same as go($\delta + \delta'$), navigate;go(-1) has the same fully active documents as doing nothing, session history can be implemented effeciently in memory...]

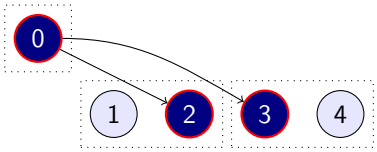
[I suspect none of these are true of the current spec, can we find a model in which they are true?]

Goal 1: If H traverses the history by δ to H' and H' traverses the history by δ' to H'' then H traverses the history by $\delta + \delta'$ to H''.

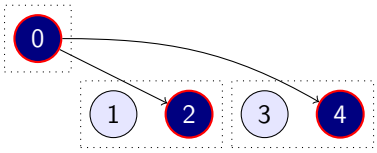
Counterexample 1: Let H be:



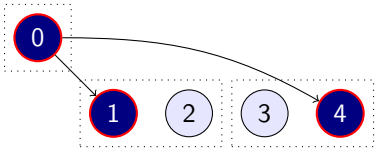
which traverses the history by 1 to:



which traverses the history by 1 to:



but H traverses the history 2 to:



This counterexample is caused by the definition of 'traverses the history by δ' which only traverses one document's session history. Instead, we should traverse the history of all δ documents.

Patch 1: Define H *traverses the history by* $+\delta$ *to* H' when:

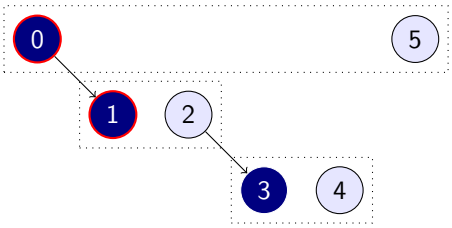
- the joint session future of H is $d_1 < \cdots < d_\delta < \cdots$,
- there is some $H = H_0, \ldots, H_\delta = H'$, such that
- H_{i-1} traverses the history to d_i in H_i for each $1 \leq i \leq \delta$.

Define H *traverses the history by* $-\delta$ *to* H' when:

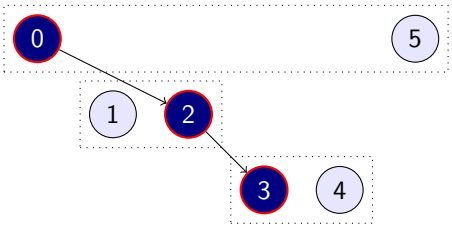
- the joint session past of H is $d_1 > \cdots > d_\delta > \cdots$,
- there is some $H = H_0, \ldots, H_\delta = H'$, such that
- H_{i-1} traverses the history to d_i in H_i for each $1 \leq i \leq \delta$.

Unfortunately, Goal 1 is not satisfied, even with this patch.

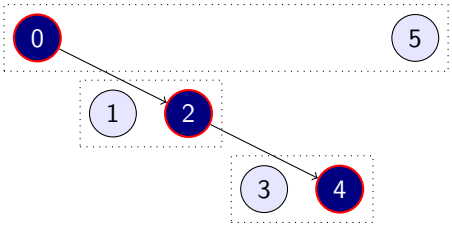
Counterexample 2: Let H be:



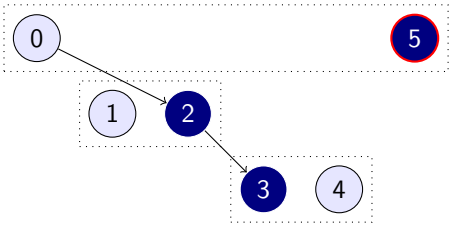
which moves forwards by 1 to:



which in turn moves forwards by 1 to:



but H goes forward by 2 to:

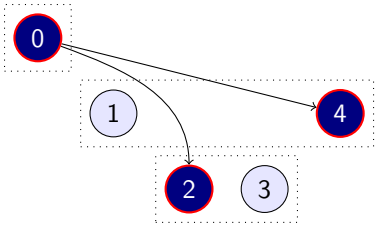


The problem this time is that the definition of 'joint session history' only includes the fully active documents, not all active documents.

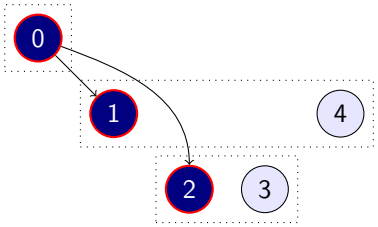
Patch 2: Define:

- the *joint session future* is $\{e \mid \exists d \in A \,.\, d \lesssim e\}$, and
- the *joint session past* is $\{e \mid \exists d \in A \,.\, d \gtrsim e\}$.

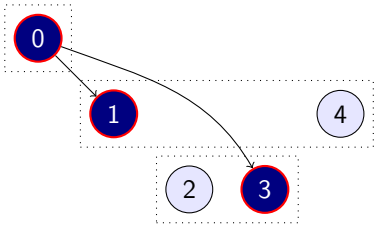
Counterexample 3: Let H be:



which traverses the history by -1 to:



which traverses the history by 1 to:



which is not the same as H.

[ASAJ: Not sure about this...]

Patch 3: Define H *traverses the history from* d' when there is some d such that:

- $d \lesssim d'$,
- for any $e \lesssim d'$ we have $e \leq d$, and
- H traverses the history to d.

Define H *traverses the history by* $-\delta$ *to* H' when:

- the joint session past and active documents of H is $d_1 > \cdots > d_\delta > \cdots$,
- there is some $H = H_0, \ldots, H_\delta = H'$, such that
- H_{i-1} traverses the history from d_i in H_i for each $1 \leq i \leq \delta$.

Goal 2: If d in H navigates to d' in H', and H' traverses the history by -1 to H'', then $FA = FA''$.

5. Experiments

[In this section various different navigation and traversal scenarios are tested in popular web browsers to see where they differ in behaviour from both the spec and each other.]

Experiment 1: In this experiment Goal 1 is tested.

- *H traverses the history by -4 to H'*
- *H' traverses the history by $+4$ to H''*

By Goal 1, these traversals should be the same thing as *H traversing by* 0 which is a no-op; therefore, $H = H''$.

Firefox:

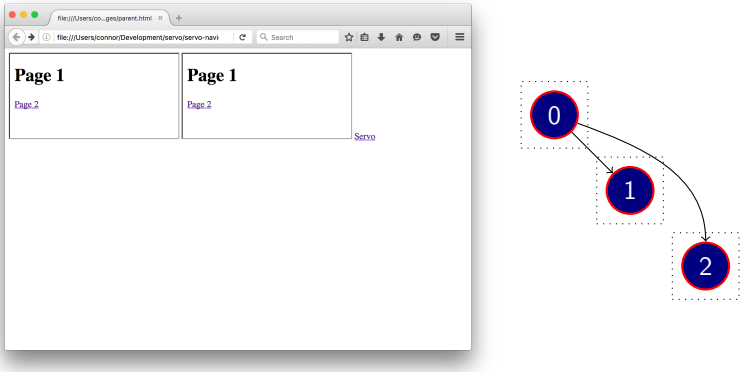


Figure 1: Initial State

Navigate document 1 to Page 2:

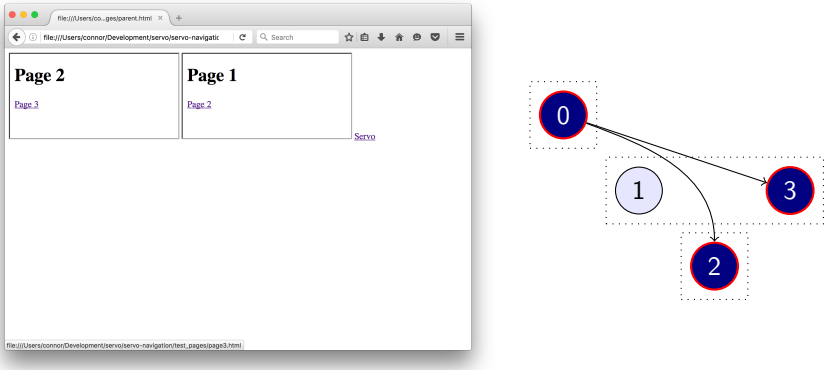


Figure 2: Navigate document 1 to Page 2.

Navigate document 3 to Page 3:

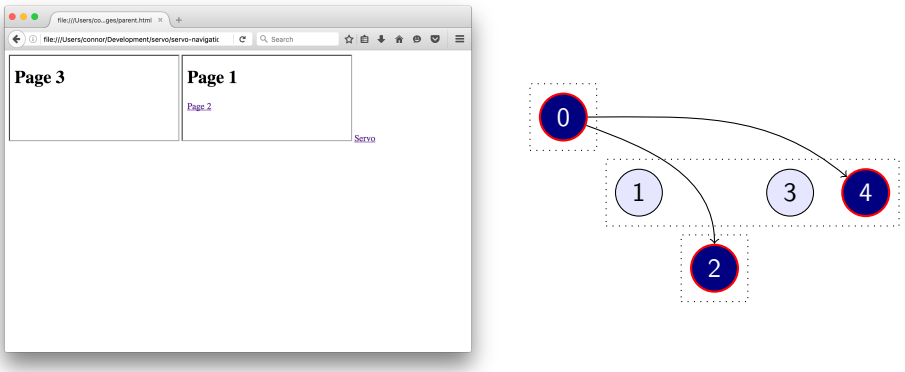


Figure 3: Navigate document 3 to Page 3.

Navigate document 2 to Page 2:

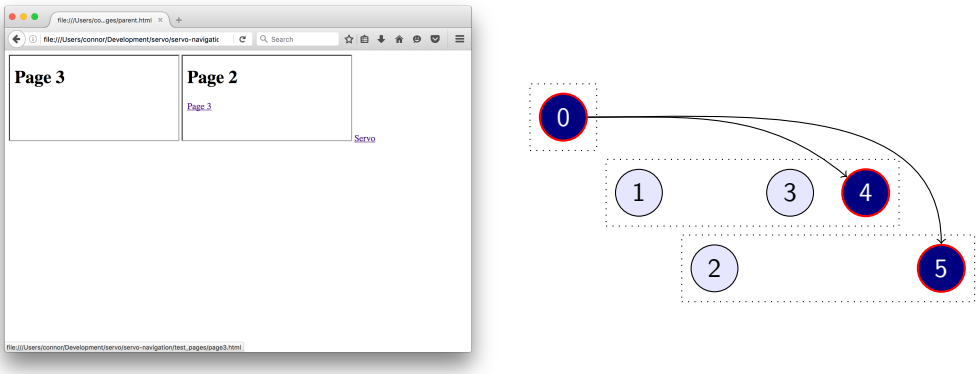


Figure 4: Navigate document 2 to Page 2.

Navigate document 5 to Page 3:

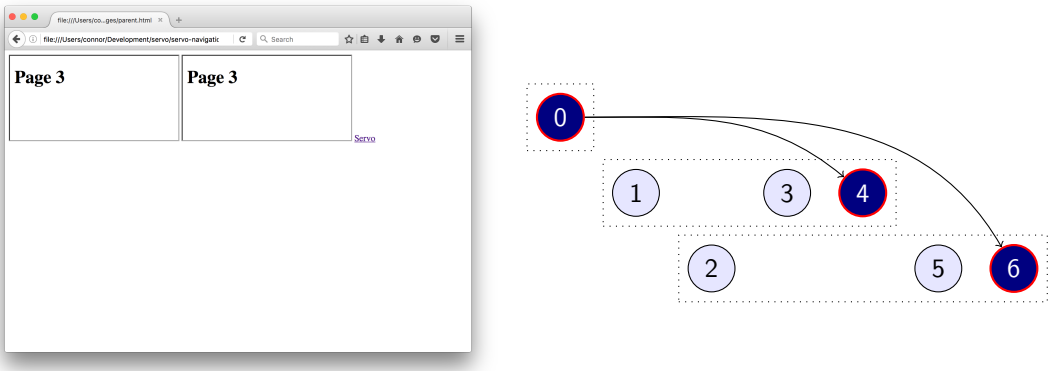


Figure 5: Navigate document 5 to Page 3.

H traverses the history by −4 to H′:

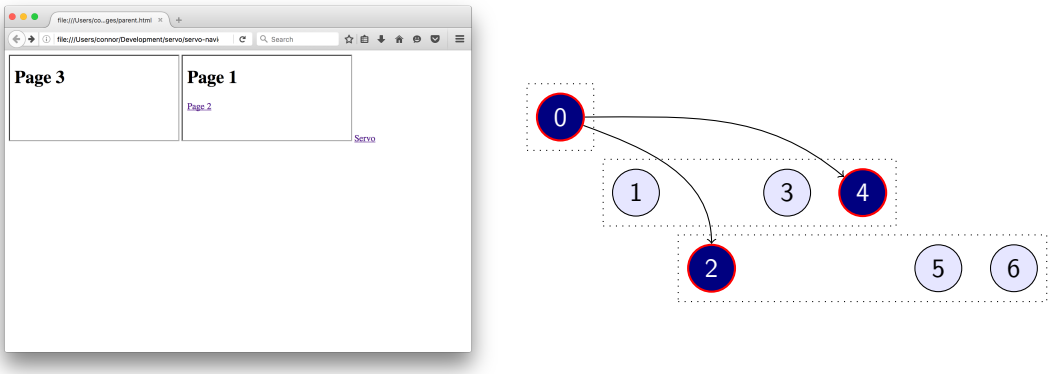


Figure 6: Traversal by -4.

This state is obviously wrong, as document 4 should have traversed to document 1. This is similar to counterexample 1.

H' traverses the history by 4 to H'':

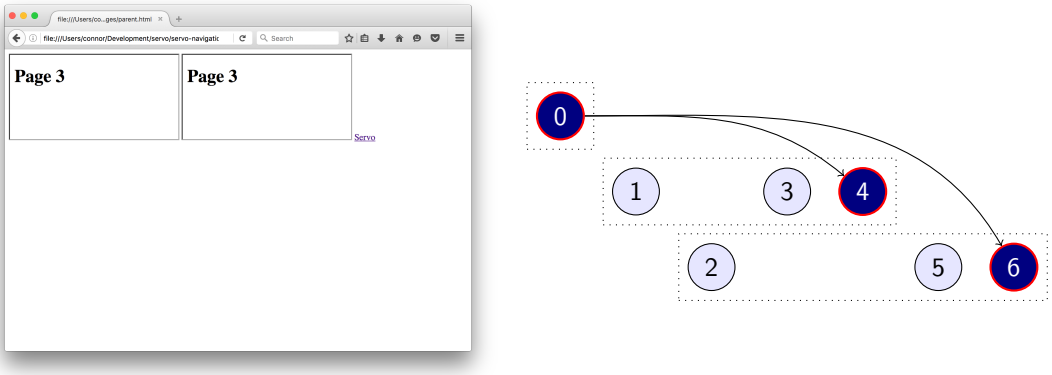


Figure 7: Traversal by 4.

While this result does satisfy Goal 1, there are still some issues:

- Figure 6 yields an incorrect traversal. [CGB: I believe this actually does break Goal 1 as navigating by -1 four times should yield the correct state.]
- It is impossible to get back to Page 1 on both Frames. [CGB: Looks to be a bug in FF, when holding down the back button, the list of pages to traverse to shows up. Clicking on the oldest item on the list does nothing and does not activate that item.]

Safari:

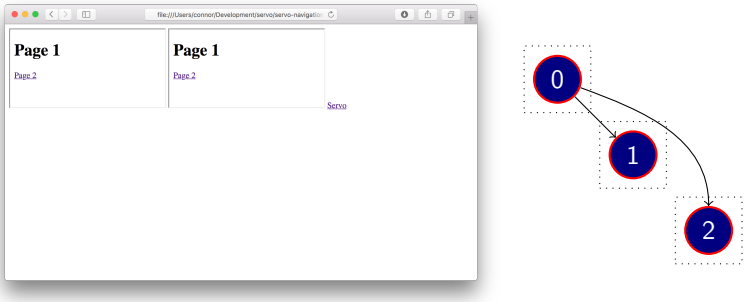


Figure 8: Initial State

Navigate document 1 to Page 2:

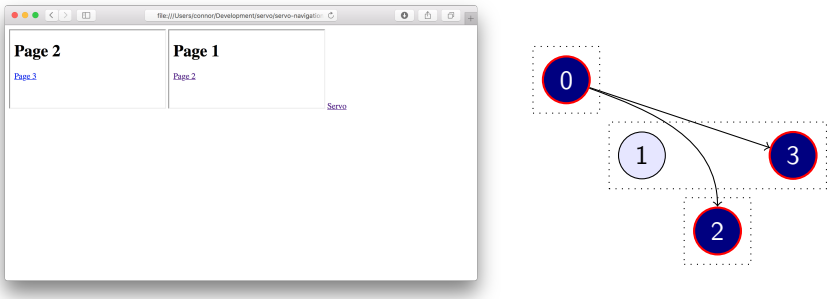


Figure 9: Navigate document 1 to Page 2.

Navigate document 3 to Page 3:

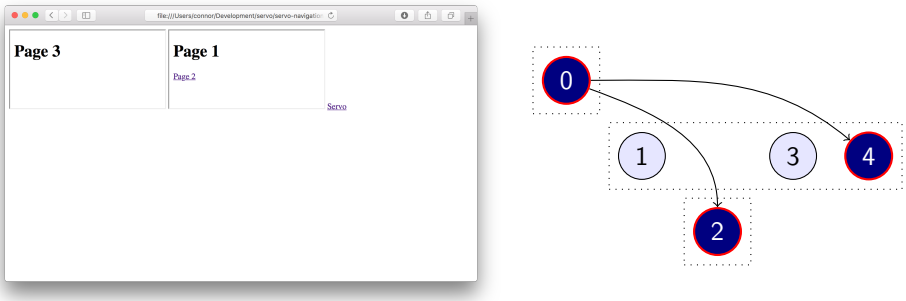


Figure 10: Navigate document 3 to Page 3.

Navigate document 2 to Page 2:

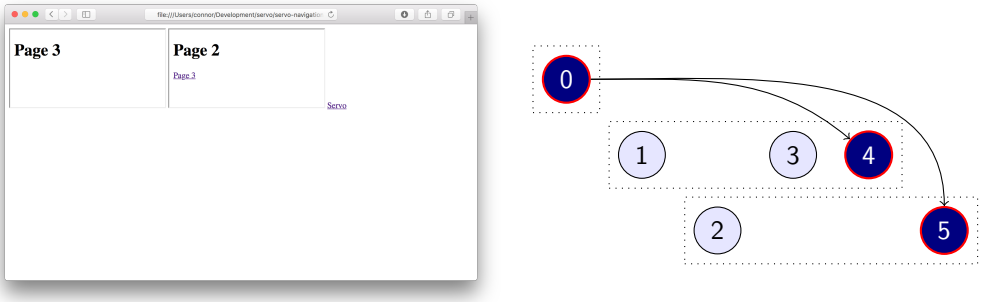


Figure 11: Navigate document 2 to Page 2.

Navigate document 5 to Page 3:

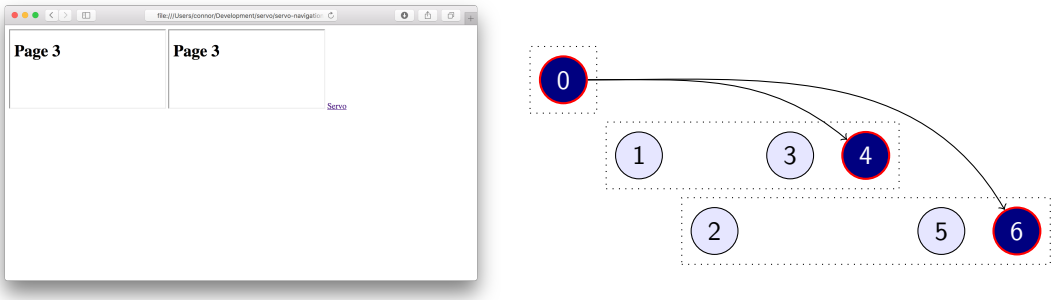


Figure 12: Navigate document 5 to Page 3.

H traverses the history by -4 *to* H':

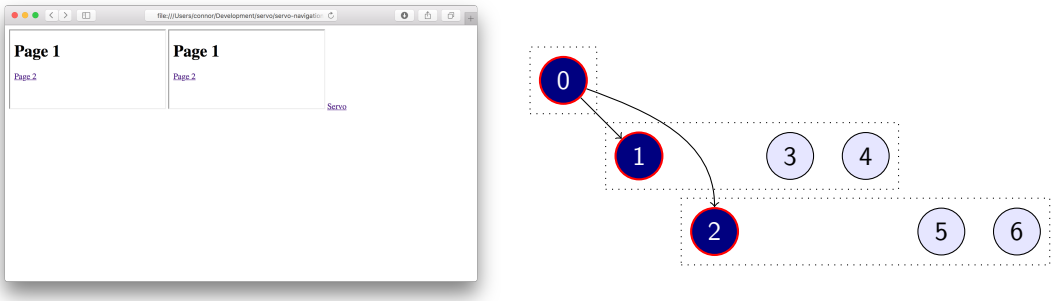


Figure 13: Traversal by -4.

H' *traverses the history by* 4 *to* H'':

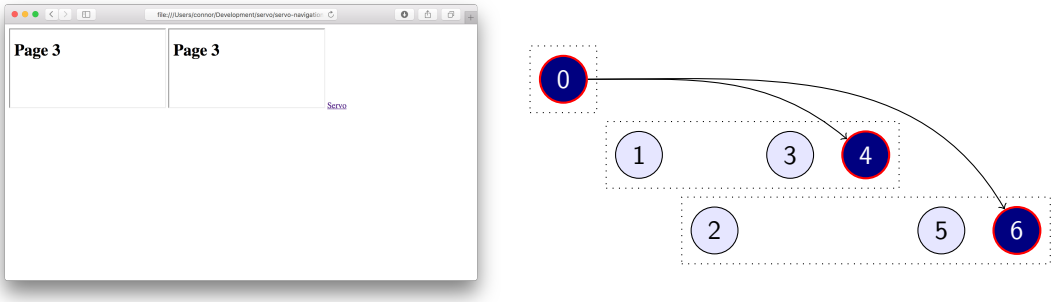


Figure 14: Traversal by 4.

These results in Safari satisfy Goal 1.
Chrome:

Figure 15: Initial State

Navigate document 1 to Page 2:

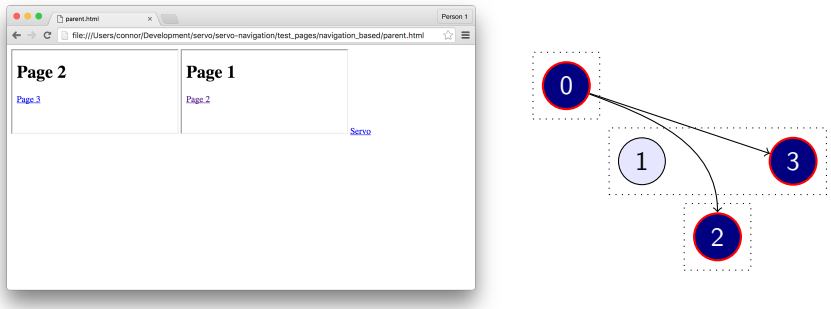


Figure 16: Navigate document 1 to Page 2.

Navigate document 3 to Page 3:

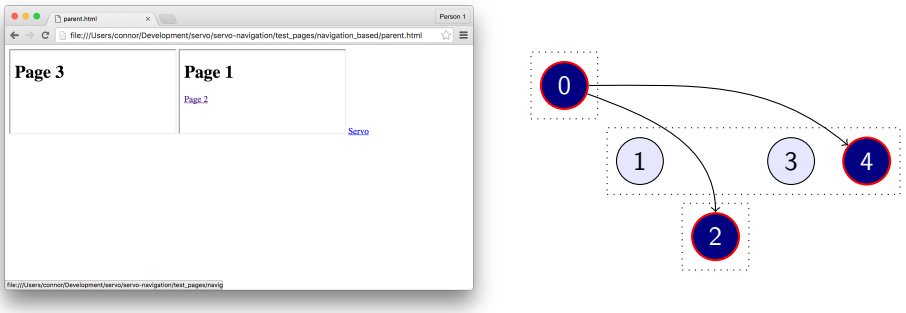


Figure 17: Navigate document 3 to Page 3.

Navigate document 2 to Page 2:

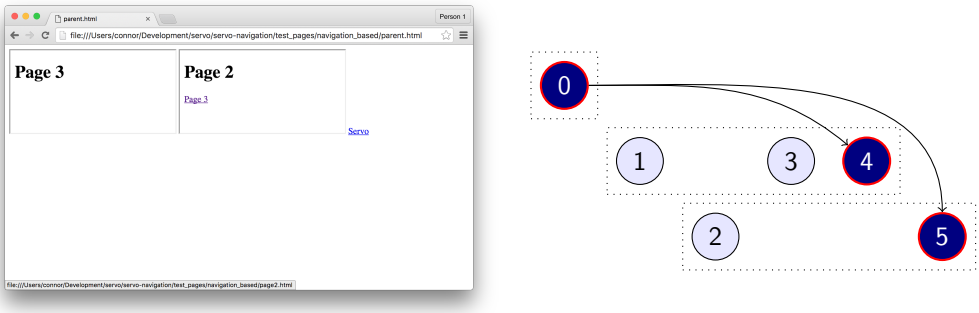


Figure 18: Navigate document 2 to Page 2.

Navigate document 5 to Page 3:

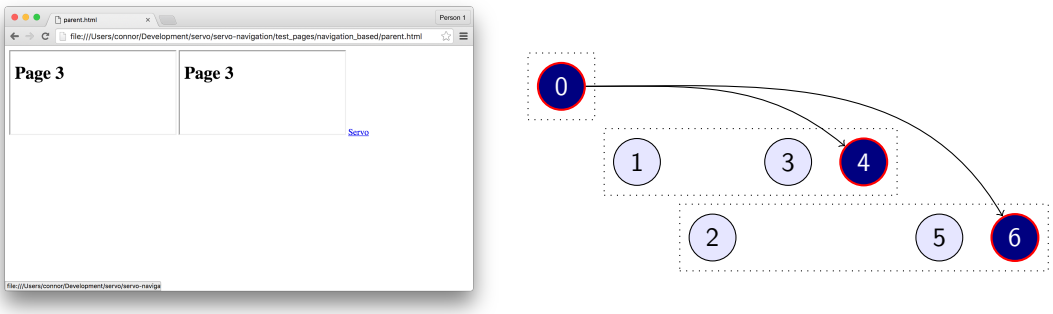


Figure 19: Navigate document 5 to Page 3.

H traverses the history by -4 *to* H':

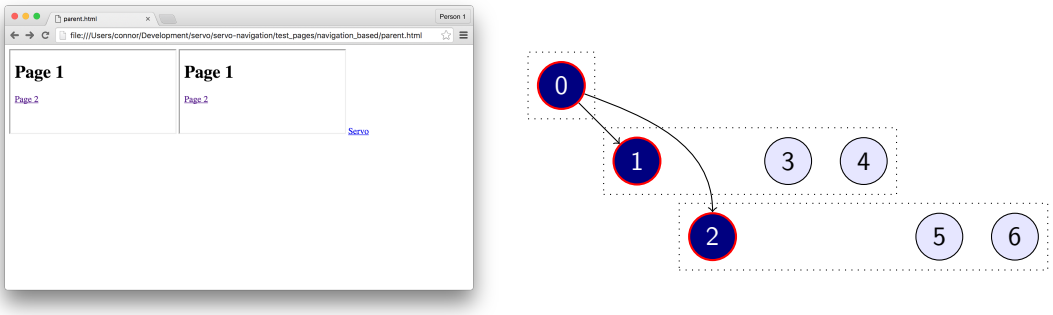


Figure 20: Traversal by -4.

H' traverses the history by 4 to H'':

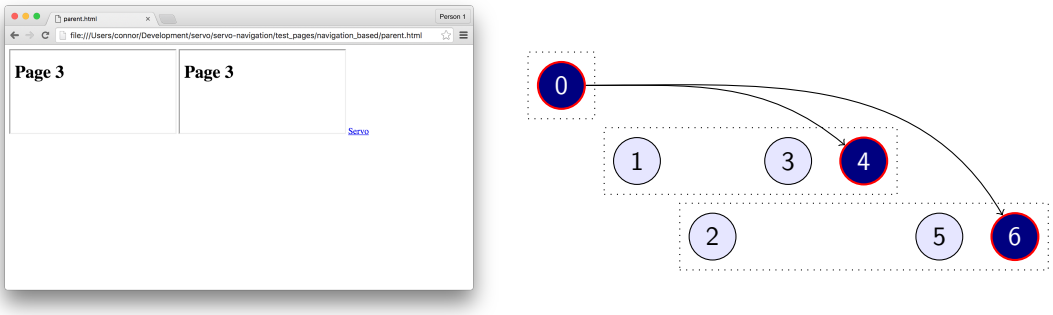


Figure 21: Traversal by 4.

These results in Chrome satisfy Goal 1.

Experiment 2: In this experiment *pushState* and *replaceState* traversals are tested against Goal 1.

Firefox:

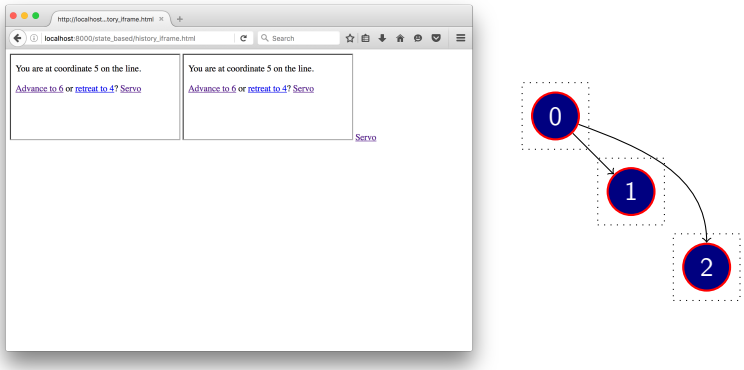


Figure 22: Initial State

Advance document 1 to 6:

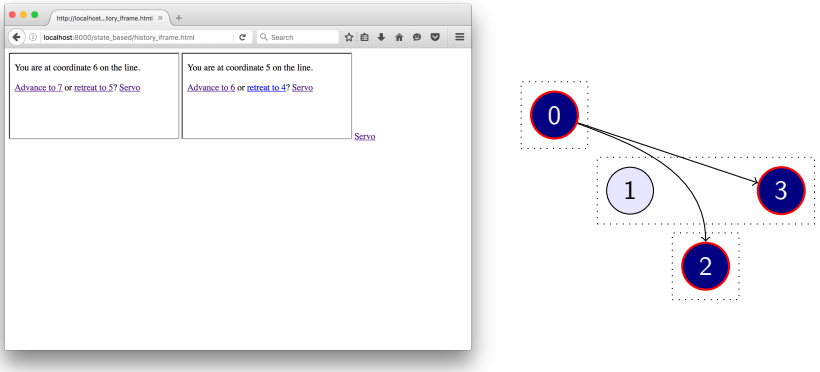


Figure 23: Advance document 1 to 6

Advance document 3 to 7:

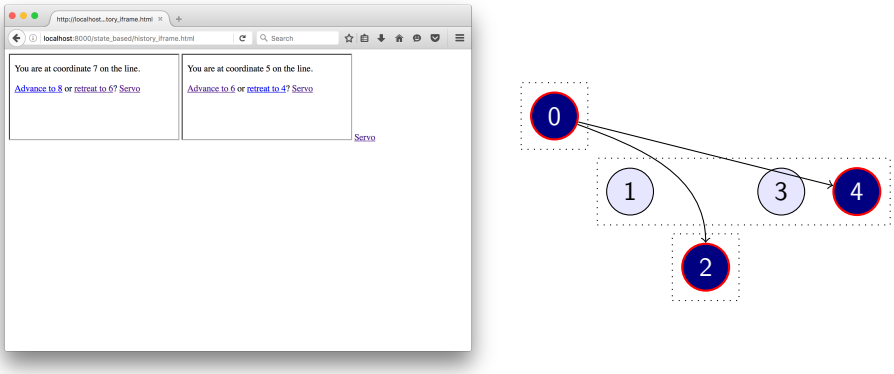


Figure 24: Advance document 3 to 7

Advance document 2 to 6:

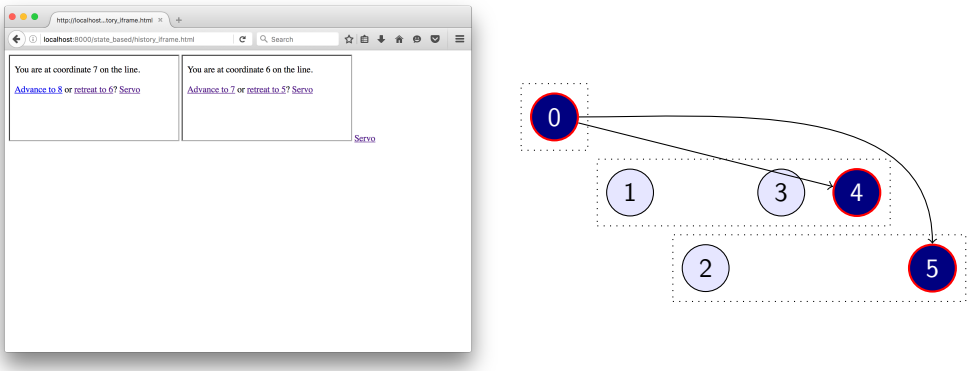


Figure 25: Advance *document*2 to 6

Advance *document*2 to 7:

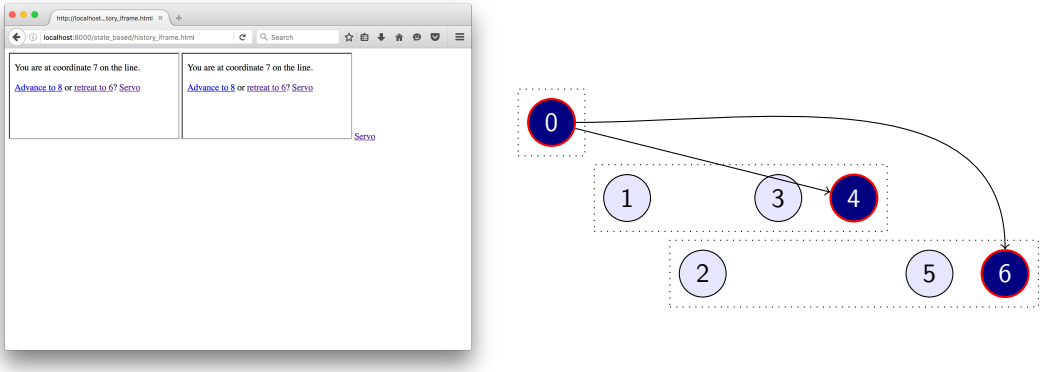


Figure 26: Advance document 2 to 7

Traverse H by -4:

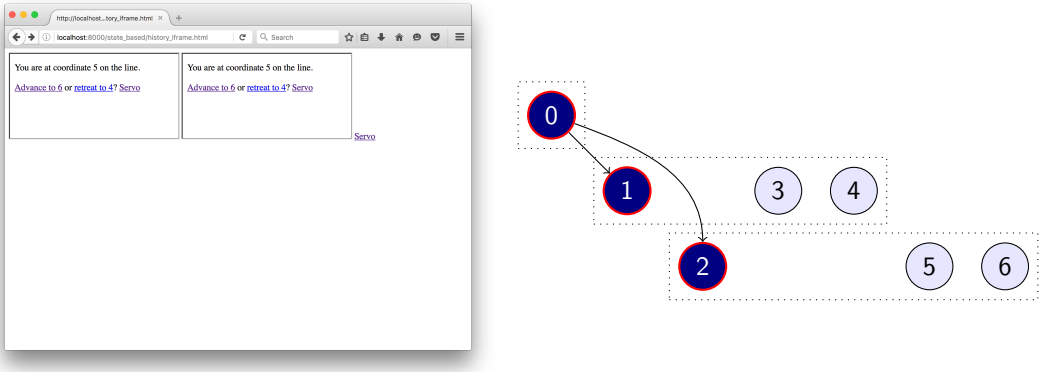


Figure 27: Traverse H by -4

Traverse H by 4:

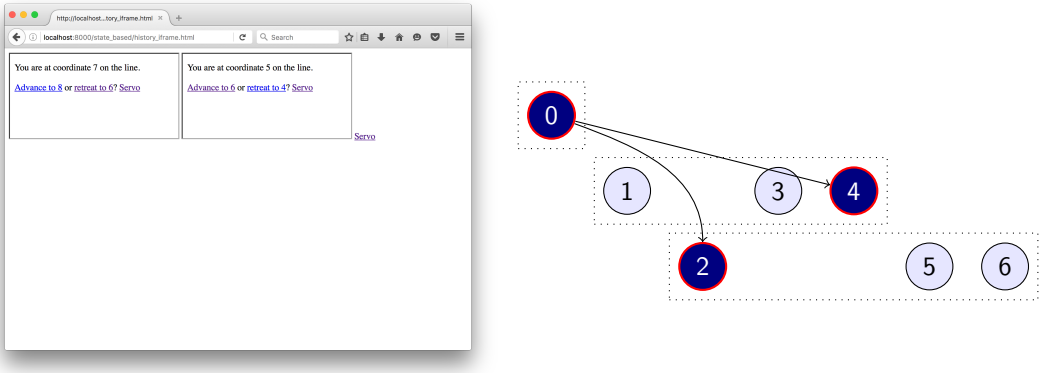


Figure 28: Traverse H by 4

The last traversal does not satisfy Goal 1.

6. Specification

[Suggested edits to the spec: 1. traverse to each document, not just the selected one, 2. keep all documents in the seession history, not just the fully active ones, 3. change the session history order.]

7. Conclusion

[We did stuff.]