Technical Report CS73004-R

A Topological Model of the Data Space for a Block Structured Language

by

T. C. Wesselkamper

Virginia Polytechnical Institute and State University

Author's address: T. C. Wesselkamper, Dept. of Computer Science, Virginia Polytechnical Institute and State University, Blacksburg, Virginia 24061

ABSTRACT

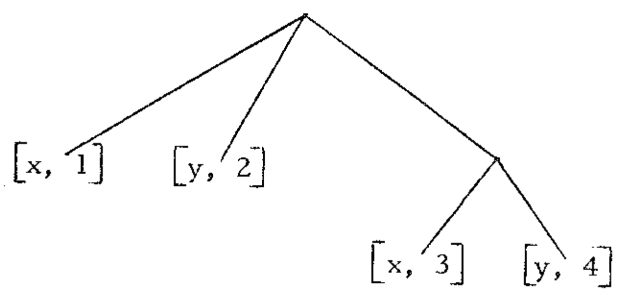
A space X is defined, the points of which are trees which possess a unique name property. A topology is defined for X based upon partial order. The topology is shown to be reasonably natural relative to the rationals. The topological space is shown to be neither Hausdorff nor T_1. The implications of this for program convergence are discussed with examples.

1. Definitions

It is natural to represent the state defined by a block structured language by a tree whose leaves are [name, value] pairs. For example, consider the Algol fragment:

```
begin integer x, y;

    x := 1;

    y := 2;

    begin integer x, y;

        x := 3;
        y := 4;.
```

The state defined by this fragment may be modelled by the tree:



Formally a tree with root x is a finite graph such that each point except x is the end point of exactly one edge, x is the endpoint of no edge, and the graph contains no cycles. [1] Contrary to general usage we consider a single leaf to be a tree.

A tree A is a maximal subtree of a tree B if there exists no tree C such that A is a proper subtree of C and C is a proper subtree is B. If A is a maximal subtree of B we write $A \in B$.

Let Q denote the field of rationals and Z the ring of integers. Let L be a denumerable set such that $t, f, u \in L$. Although Q contains an isomorphic copy of Z, $Q \cap Z = \emptyset$. Let $L \cap Q = L \cap Z = \emptyset$. Let $V = Q \cup Z \cup L$. Let N be a denumerable set of names such that $N \cap V = \emptyset$. The Cartesian product $N \times V$ forms the set of leaves of the trees we define.

If $\{x_1, x_2, \ldots, x_n\}$ is the set of maximal subtrees of a tree A, we write $A = (x_1, x_2, \ldots, x_n)$, which is less graphic but more tractable typographically.

A tree A with leaves in $N \times V$ possesses the unique name property if for each subtree B of A and each pair of leaves $[x, x']$ and $[y, y']$ of B, $x = y$ implies that $x' = y'$. If a tree possesses the unique name property, it is clear that each subtree possesses the unique name property.

We define the data space X as follows:

1.) if $x \in N \times V$, then $x \in X$;

2.) if $x_1, x_2, \ldots, x_n \in X$ and if $(x_1, x_2, \ldots, x_n)$ possesses the unique name property, then

$$(x_1, x_2, \ldots x_n) \in X.$$

This same construction has been studied by the author elsewhere in connection with programming languages. [3]

2. A Topology for X.

Throughout this paper we use the definitions, notation, and theorems of Kelley[2].

If x and y are elements of X, say that xRy if $x = y$ or there exists a set $\{x, a_1, \ldots, a_n, y\}$ (called a chain) such that $x \in a_1 \in a_2 \in \ldots \in a_n \in y$. It is trivially true that the relation R is reflexive and transitive. That R is anti-symmetric follows from that fact that no set is an element of itself.

Let I denote a non-void open interval of the rational line with the usual topology.

For each $n \in N$ let $nI = \{[n, i] \; : i \in I\}$, a set of points in $N \times V$.

For $x \in X$ define:

$$O_x = \{z: \; xRz\} .$$

For $n \in N$ and $I \subset Q$, define:

$$O_{nI} = \bigcup_{x \in nI} O_x$$

Finally define:

$$\mathcal{S} = \{O_{nI}: \; n \in N, \; I \subset Q\} \cup \{O_x: \; x = [n, v] \; , n \in N, \; v \in Z \cup L\}$$

This family $\mathcal{S}$ is a subbase for a topology $\mathcal{T}$ for X. Since $Z \cup L$ is countable and since the set of intervals of Q is countable, the family $\mathcal{S}$ is countable. A base for the topology $\mathcal{T}$ is the family of finite intersections of members of $\mathcal{S}$. Hence $\mathcal{T}$ has a countable base and so is said to satisfy "the second axiom of countability." [2, p. 48]

<u>Lemma 1</u> -- Any two sets O_x and O_y have a non-void intersection.

<u>proof</u> -- Let z = (x, (y)). Then $x \in Z$ and $y \in (y) \in z$. Hence both xRz and yRz, that is $z \in O_x$ and $z \in O_y$, and so $z \in O_x \cap O_y$. Note that is both x and y are elements of N x V it is possible that (x, y) may not satisfy the unique name condition and so (x, y) may not be an element of X; (x, (y)) is certainly in X.

<u>Corollary</u> -- Each two members of $\mathcal{S}$ have a non-void intersection.

<u>proof</u> -- Each O_{nI} is a union of sets O_x. If the two members of $\mathcal{S}$ are of the form O_{nI} and O_{n*I*}, then choose $O_x \subset O_{nI}$ and $O_{x*} \subset O_{n*I*}$, O_x and O_{x*} have a non-void intersection, and hence so do O_{nI} and O_{n*I*}.

Lemma 2 -- If U is an open set and $x \in U$, then $O_x \subset U$.

proof -- If U is open then $U = \bigcup U_\alpha$, for α in some index set A, where each U_α is a finite intersection of members of $\mathcal{S}$. If $x \in U$, then $x \in U_a$ and some element a of A. There is a finite set $B \subset N \times V$ such that $U_a = \bigcap_{y \in B} O_y$. Thus $x \in O_y$ for each $y \in B$. Hence yRx for each $y \in B$. Let w be an arbitrary point of O_x. Then xRw and by the transitivity of R, yRw for each $y \in B$, that is, $w \in O_y$ for each y in B, and so $w \in U_a$. Since w is an arbitrary point of O_x, we have $O_x \subset U_a \subset U$.

Lemma 3 -- There are points of X which are not closed.

proof -- Denote the complement of x in X by Cx. If $a \in X$, then $(a) \in X$. The point (a) is closed if and only if C(a) is open. Since $a \neq (a)$, $a \in C(a)$. Suppose that C(a) is open. Then by Lemma 2, $a \in O_a \subset C(a)$. But $(a) \in O_a$ since aR(a). This is a contradiction. Thus C(a) is not open and (a) is not closed.

Corollary -- The space X is neither a T_1 space nor a Hausdorff space.

proof -- A space of T_1 if its points are closed. A necessary condition that a space be Hausdorff is that it be T_1. [2, p. 56] That X is not a Hausdorff space could be proved directly from the fact that no two open sets of X are disjoint.

It is easy, but not important here, to prove that X is a T_0 space.

A space which satisfies the second axiom of countability is Hausdorff if and only if each sequence in the space converges to at most one point. [2, pp. 67 and 72]. The space X satisfies the second axiom of countability and is not Hausdorff. Hence there are sequences in the space which converge to more than one point. We consider these in the next section.

A space is called countably compact if each countable open cover has a finite subcover. A space is called sequentially compact if each sequence in the space has a convergent subsequence. A space which satisfies the second axiom of countability is countably compact if and only if it is sequentially compact. [2, pp. 162, 50]

Lemma 4 -- The space X is not countably compact.

proof -- The family of sets $\mathcal{S}$ is a countable open cover for X. If $n \in N$, and $t \in L$, then, $[n, t] = y \in N \times V$, and so $O_y \in \mathcal{S}$. Suppose that there is another O_x in $\mathcal{S}$ such that $y \in O_x$. Then xRy and since y is a leaf there is no chain from x to y and so $x = y$. Any finite subfamily of the cover $\mathcal{S}$ omits points of $(Z \cup L)$ and so is not a subcover.

Corollary -- There are sequences in X which do not converge.

Examples of such sequences will be seen in the next section.

Corollary -- The space X is not compact.

3. The Convergence of Programs.

It is convenient to treat every sequence of points as if it were an infinite sequence. In the usual way we say that if S(n) is a finite sequence and if S(N) is the last point of the sequence, then we define $S(i) = S(N)$ for all $i > N$.

In the last section we noted that a sequence may converge to more than one point in X. This behavior is exhibited by the Algol program:

```
begin integer x, y, z;
        x := 1;
        y := 2;
        z := 3;
L:      z := z + 1;
        goto L
end
```

If we adopt the convention that at the time of declaration an identifier is initialized to the value u, then the assignment statements define a sequence of points:

$$([x, u], \quad [y, u], \quad [z, u])$$

$$([x, 1], \quad [y, u], \quad [z, u])$$

$$([x, 1], \quad [y, 2], \quad [z, 3])$$

$$([x, 1], \quad [y, 2], \quad [z, 4])$$

etc.

This sequence of points is eventually in both the sets $O_{[x, 1]}$ and $O_{[y, 2]}$ and thus the sequence converges to both the points $[x, 1]$ and $[y, 2]$. Were x, y, z to have been <u>real</u> in the program, the integers 1, 2, 3, etc. would have been replaced by their isomorphic images 1.0, 2.0, 3.0, etc. Then for each open interval I_1 about 1.0 and each open interval I_2 about 2.0, the sequence would eventually be in O_{xI_1} and O_{yI_2}.

It should be noted that if E is the equivalence relation defined over N x Q by $[x, x'] \; E \; [y, y']$ if and only if $x' = y'$, and if the topology $\mathcal{T}$ is relativized to N x Q, then the quotient topology (N x Q)/E is the usual topology for the rationals. As a result,

References

1. C. Berge, _The Theory of Graphs and Its Applications_, (New York: Wiley, 1962), p. 160.

2. John L. Kelley, _General Topology_, (New York: Van Nostrand, 1955)

3. T.C. Wesselkamper and Eric Nixon, "A Complete Horizontal Microlanguage", (submitted to _JACM_)