Consider an XML document containing medical data as in Figure 3. Each node describes a person with an attribute assigning a name and an attribute stating whether or not the person has or had leukemia. The child-of relation in the tree models the real child-of relation between persons. Thus the root of the tree is person `a` which has leukemia. Person `a` has a child `a1` without the disease and a grandchild `a12` which has it.

Consider the following information need: *find persons for whom all its ancestors*[4] *had/have leukemia, but the person itself has not*. The answer set of this query in the example document is the set of nodes with name attribute `a1` and `a22`. The information need can be expressed in first order logic using a suitable signature. Let `child` and `descendant` be binary and `P` and `has_leukemia` be unary predicates. The information need is expressed by a first order formula in one free variable:

$$\begin{aligned}\mathtt{P}(x) \wedge \neg\mathtt{has\_leukemia} \wedge \quad & x = \mathit{root} \ \vee \\ & \mathtt{has\_leukemia}(\mathit{root}) \wedge \mathit{root}\,\mathtt{child}\,x \ \vee \\ & \forall z((\mathtt{descendant}(\mathit{root}, z) \wedge \mathtt{descendant}(z, x)) \rightarrow \mathtt{has\_leukemia}(z)).\end{aligned}$$

This information need can be expressed in XPath 1.0 for arbitrarily deep documents by the infinite disjunction

```
/P[@leukemia='no'] |
  /P[@leukemia='yes']/P[@leukemia='no'] |
    /P[@leukemia='yes']/P[@leukemia='yes']/P[@leukemia='no'] ...
```

but it is not possible to express this information need in XPath 1.0 (at least not in Core XPath as defined by Gottlob et al, the proof uses the fact that the binary “until” operator is not definable in temporal logic with only the unary “sometimes in the future” and “sometimes in the past”) . What seems to be needed is the notion of a *conditional path*. Let $_{[\mathtt{test}]}\mathtt{child}$ denote all pairs (n, n') such that n' is a child of n and the `test` succeeds at n. Then the information need can be expressed by

$$/({}_{[\mathtt{@leukemia='yes'}]}\mathtt{child})^* :: \mathtt{P}[\mathtt{@leukemia} =' \mathtt{no}'].$$

The axis $({}_{[\mathtt{@leukemia='yes'}]}\mathtt{child})^+$ describes a path $n_0, n_1, \ldots, n_k$, for $k \geq 1$ such that at all nodes $n_0, \ldots n_{k-1}$ the leukemia attribute equals `yes`.

```
<P name=a leukemia=yes>
   <P name=a1 leukemia=no>
       <P name=a11 leukemia=no/>
       <P name=a12 leukemia=yes/>
       <P name=a13 leukemia=no/>
   </P>
   <P name=a2 leukemia=yes>
       <P name=a21 leukemia=yes/>
       <P name=a22 leukemia=no/>
   </P>
</P>
```

Figure 3: XML document containing medical data.

[4]Obviously persons have two parents. In the example, we assume that the ancestors of a person are just the nodes reachable by traveling upward in the tree.