Transitive closure according to Roy-Floyd-Warshall

Makarius Wenzel

August 16, 2018

Abstract

This formulation of the Roy-Floyd-Warshall algorithm for the transitive closure bypasses matrices and arrays, but uses a more direct mathematical model with adjacency functions for immediate predecessors and successors. This can be implemented efficiently in functional programming languages and is particularly adequate for sparse relations.

Contents

1 Transitive closure algorithm

2 Correctness proof
   2.1 Miscellaneous lemmas
   2.2 Bounded closure
   2.3 Main theorem

3 Alternative formulation

1 Transitive closure algorithm

The Roy-Floyd-Warshall algorithm takes a finite relation as input and produces its transitive closure as output. It iterates over all elements of the field of the relation and maintains a cumulative approximation of the result: step 0 starts with the original relation, and step $\text{Suc } n$ connects all paths over the intermediate element $n$. The final approximation coincides with the full transitive closure.

This algorithm is often named after “Floyd”, “Warshall”, or “Floyd-Warshall”, but the earliest known description is due to B. Roy [1].

Subsequently we use a direct mathematical model of the relation, bypassing matrices and arrays that are usually seen in the literature. This is more efficient for sparse relations: only the adjacency for immediate predecessors and successors needs to be maintained, not the square of all possible
combinations. Moreover we do not have to worry about mutable data structures in a multi-threaded environment. See also the graph implementation in the Isabelle sources $\texttt{ISABELLE_HOME/src/Pure/General/graph.ML}$ and $\texttt{ISABELLE_HOME/src/Pure/General/graph.scala}$.

type-synonym relation = (\texttt{nat} \times \texttt{nat}) \texttt{set}

fun steps :: relation $\Rightarrow$ nat $\Rightarrow$ relation
where
  steps rel 0 = rel
| steps rel (Suc n) =
    steps rel n $\cup$ \{(x, y). (x, n) $\in$ steps rel n $\land$ (n, y) $\in$ steps rel n\}

Implementation view on the relation:

definition preds :: relation $\Rightarrow$ nat $\Rightarrow$ nat set
where
  preds rel y = \{x. (x, y) $\in$ rel\}

definition succs :: relation $\Rightarrow$ nat $\Rightarrow$ nat set
where
  succs rel x = \{y. (x, y) $\in$ rel\}

lemma
  steps rel (Suc n) =
  steps rel n $\cup$ \{(x, y). x $\in$ preds (steps rel n) n $\land$ y $\in$ succs (steps rel n) n\}
  by (simp add: preds-def succs-def)

The main function requires an upper bound for the iteration, which is left unspecified here (via Hilbert’s choice).

definition is-bound :: relation $\Rightarrow$ nat $\Rightarrow$ bool
where
  is-bound rel n $\leftarrow\rightarrow$ (\forall m $\in$ Field rel. m < n)

definition transitive-closure rel = steps rel (SOME n. is-bound rel n)

2 Correctness proof

2.1 Miscellaneous lemmas

lemma finite-bound:
  assumes finite rel
  shows $\exists$ n. is-bound rel n
  using assms
proof induct
  case empty
  then show ?case by (simp add: is-bound-def)
next
  case (insert p rel)
  then obtain n where n: \forall m $\in$ Field rel. m < n
    unfolding is-bound-def by blast
  obtain x y where p = (x, y) by (cases p)
  then have \forall m $\in$ Field (insert p rel). m < max (Suc x) (max (Suc y) n)
using n by auto
then show ?case
unfolding is-bound-def by blast
qed

lemma steps-Suc: (x, y) ∈ steps rel (Suc n) ↔
(x, y) ∈ steps rel n ∨ (x, n) ∈ steps rel n ∧ (n, y) ∈ steps rel n
by auto

lemma steps-cases:
assumes (x, y) ∈ steps rel (Suc n)
obtains (copy) (x, y) ∈ steps rel n
| (step) (x, n) ∈ steps rel n and (n, y) ∈ steps rel n
using assms by auto

lemma steps-rel: (x, y) ∈ rel ⇒ (x, y) ∈ steps rel n
by (induct n) auto

2.2 Bounded closure

The bounded closure connects all transitive paths over elements below a
given bound. For an upper bound of the relation, this coincides with the
full transitive closure.

inductive-set Clos :: relation ⇒ nat ⇒ relation
for rel :: relation and n :: nat
where
  base: (x, y) ∈ Clos rel n if (x, y) ∈ rel
| step: (x, y) ∈ Clos rel n if (x, z) ∈ Clos rel n and (z, y) ∈ Clos rel n and z <
n
theorem Clos-closure:
assumes is-bound rel n
shows (x, y) ∈ Clos rel n ↔ (x, y) ∈ rel+
proof
show (x, y) ∈ rel+ if (x, y) ∈ Clos rel n
using that by induct simp-all
show (x, y) ∈ Clos rel n if (x, y) ∈ rel+
using that
proof (induct rule: trancl-induct)
case (base y)
  then show ?case by (rule Clos.base)
next
case (step y z)
from (y, z) ∈ rel have 1: (y, z) ∈ Clos rel n by (rule base)
from (y, z) ∈ rel and is-bound rel n have 2: y < n
unfolding is-bound-def Field-def by blast
from step(3) 1 2 show ?case by (rule Clos.step)
qed
qed
lemma Clos-Suc:
assumes \((x, y) \in \text{Clos rel } n\)
shows \((x, y) \in \text{Clos rel } (\text{Suc } n)\)
using assms by induct (auto intro: Clos.intros)

In each step of the algorithm the approximated relation is exactly the bounded closure.

theorem steps-Clos-equiv: \((x, y) \in \text{steps rel } n \iff (x, y) \in \text{Clos rel } n\)
proof (induct \(n\) arbitrary: \(x\) \(y\))
case 0
  show ?case
  proof
    show \((x, y) \in \text{Clos rel } 0\) if \((x, y) \in \text{steps rel } 0\)
    proof
      from that have \((x, y) \in \text{rel}\) by simp
      then show \(?\text{thesis}\) by (rule Clos.base)
    qed
    show \((x, y) \in \text{steps rel } 0\) if \((x, y) \in \text{Clos rel } 0\)
    using that by cases simp-all
  qed
next
case (Suc \(n\))
  show ?case
  proof
    show \((x, y) \in \text{Clos rel } (\text{Suc } n)\) if \((x, y) \in \text{steps rel } (\text{Suc } n)\)
    using that
    proof (cases rule: steps-cases)
      case copy
      with Suc(1) have \((x, y) \in \text{Clos rel } n\) ..
      then show \(?\text{thesis}\) by (rule Clos-Suc)
    next
case step
      with Suc have \((x, n) \in \text{Clos rel } n\) and \((n, y) \in \text{Clos rel } n\)
      by simp-all
      then have \((x, n) \in \text{Clos rel } (\text{Suc } n)\) and \((n, y) \in \text{Clos rel } (\text{Suc } n)\)
      by (simp-all add: Clos-Suc)
      then show \(?\text{thesis}\) by (rule Clos.step) simp
    qed
    show \((x, y) \in \text{steps rel } (\text{Suc } n)\) if \((x, y) \in \text{Clos rel } (\text{Suc } n)\)
    using that
    proof induct
      case (base \(x\) \(y\))
      then show \(?\text{case}\) by (simp add: steps-rel)
    next
case (step \(x\) \(z\) \(y\))
    with Suc show \(?\text{case}\)
      by (auto simp add: steps-Suc less-Suc-eq intro: Clos.step)
  qed

4
2.3 Main theorem

The main theorem follows immediately from the key observations above. Note that the assumption of finiteness gives a bound for the iteration, although the details are left unspecified. A concrete implementation could choose the the maximum element + 1, or iterate directly over the data structures for the \textit{preds} and \textit{succs} implementation.

\textbf{theorem} transitive-closure-correctness:
\begin{enumerate}[itemsep=0pt, listparindent=0pt]
  \item assumes finite rel
  \item shows transitive-closure rel = rel^+
\end{enumerate}

\textbf{proof} –
\begin{enumerate}[itemsep=0pt, listparindent=0pt]
  \item let \(?N = \text{SOME } n\), is-bound rel \(n\)
  \item have is-bound: is-bound rel \(?N\)
    \begin{enumerate}[itemsep=0pt, listparindent=0pt]
      \item by (rule someI-ex) (rule finite-bound \[\text{OF } \text{finite rel}]\)
    \end{enumerate}
  \item have \((x, y) \in \text{steps rel } ?N \iff (x, y) \in \text{rel}^+ \text{ for } x y\)
  \item proof –
    \begin{enumerate}[itemsep=0pt, listparindent=0pt]
      \item have \((x, y) \in \text{steps rel } ?N \iff (x, y) \in \text{Clos rel } ?N\)
        \begin{enumerate}[itemsep=0pt, listparindent=0pt]
          \item by (rule steps-Clos-equiv)
        \end{enumerate}
      \item also have \(\ldots \iff (x, y) \in \text{rel}^+\)
        \begin{enumerate}[itemsep=0pt, listparindent=0pt]
          \item using is-bound by (rule Clos-closure)
        \end{enumerate}
      \item finally show \(?\text{thesis }\).
    \end{enumerate}
  \item qed
  \item then show \(?\text{thesis}\) unfolding transitive-closure-def by auto
\end{enumerate}
\begin{enumerate}[itemsep=0pt, listparindent=0pt]
  \item qed
\end{enumerate}

3 Alternative formulation

The core of the algorithm may be expressed more declaratively as follows, using an inductive definition to imitate a logic-program. This is equivalent to the function specification \textit{steps} from above.

\textbf{inductive} \textit{Steps} :: \textit{relation} \Rightarrow \textit{nat} \Rightarrow \textit{nat} \times \textit{nat} \Rightarrow \textit{bool}
\begin{enumerate}[itemsep=0pt, listparindent=0pt]
  \item for \textit{rel} :: \textit{relation}
\begin{enumerate}[itemsep=0pt, listparindent=0pt]
  \item base: \textit{Steps} rel 0 \((x, y)\) if \((x, y) \in \text{rel}\)
  \item copy: \textit{Steps} rel \((\text{Suc } n)\) \((x, y)\) if \textit{Steps} rel \(n\) \((x, y)\)
  \item step: \textit{Steps} rel \((\text{Suc } n)\) \((x, y)\) if \textit{Steps} rel \(n\) \((x, n)\) \textbf{and} \textit{Steps} rel \(n\) \((n, y)\)
\end{enumerate}
\end{enumerate}

\textbf{lemma} steps-equiv: \((x, y) \in \text{steps rel } n \iff \text{Steps} \text{ rel } n \ (x, y)\)

\textbf{proof} –
\begin{enumerate}[itemsep=0pt, listparindent=0pt]
  \item show \(\text{Steps} \text{ rel } n \ (x, y) \iff \text{steps} \text{ rel } n \ (x, y)\)
    \begin{enumerate}[itemsep=0pt, listparindent=0pt]
      \item using that
    \end{enumerate}
  \item proof (induct \(n\) arbitrary: \(x y\))
    \begin{enumerate}[itemsep=0pt, listparindent=0pt]
      \item case 0
        \begin{enumerate}[itemsep=0pt, listparindent=0pt]
          \item then have \((x, y) \in \text{rel}\) by simp
        \end{enumerate}
    \end{enumerate}
\end{enumerate}
then show ?case by (rule base)

next
case (Suc n)
from Suc(2) show ?case
proof (cases rule: steps-cases)
case copy
  with Suc(1) have Steps rel n (x, y).
then show ?thesis by (rule Steps.copy)
next
case step
  with Suc(1) have Steps rel n (x, n) and Steps rel n (n, y)
  by simp-all
then show ?thesis by (rule Steps.step)
qed
qed
show (x, y) ∈ steps rel n if Steps rel n (x, y)
using that by induct simp-all
qed

References