

# Optimierungen von Tarjans SZK-Algorithmus

Christian Koch

Institut für Informatik, Universität Rostock, D-18051 Rostock, Germany.

**Zusammenfassung** Dieses Dokument umfasst alle im Rahmen der Implementierung im Projekt LoLA vorgenommenen Optimierungen des Tarjan-Algorithmus mit (teils informalen) Beweisen. Es wird die Kenntnis des originalen Tarjan-Algorithmus, sowie der üblichen Notation für tiefensuchbasierende Algorithmen vorausgesetzt, wie etwa die Farbkodierung der Knoten.

## 1 Tarjan-Stack

Der Tarjan-Algorithmus verwaltet zwei Stacks, den DFS-Stack und den Tarjan-Stack. Der Tarjan-Stack enthält dabei stets die Elemente des DFS-Stack, und alle schwarzen Knoten, die in noch nicht abgeschlossenen Zusammenhangskomponenten liegen.

In der LoLA-Implementation wird der Tarjan-Stack nun derart verändert, dass zu jedem Zeitpunkt nur die nicht auf dem DFS-Stack liegenden Knoten abgelegt sind, um auf diese Weise die doppelte Speicherung zu vermeiden und somit den Speicherbedarf zu verringern. Dazu werden die Knoten nicht beim Betreten auf den Tarjan-Stack gelegt, sondern erst beim Verlassen des Knotens. Es ist zu beachten, dass diese Änderung die Reihenfolge, in der die Elemente auf dem Tarjan-Stack liegen, verändert. Insbesondere liegt der Knoten, mit dem die SZK betreten wurde und bei der sie schließlich abgeschlossen wird, nicht auf dem Tarjan-Stack, um die gefundene SZK von den restlichen Knoten weiter unten auf dem Stack abzugrenzen. Stattdessen wird die DFS-Nummer verwendet: Alle Elemente, die eine größere DFS-Nummer als der Startknoten der SZK haben, gehören dazu und werden abgeräumt, alle darunter liegenden Knoten mit kleineren DFS-Nummern gehören zu anderen SZKs und bleiben auf dem Stack.

Es ist zu zeigen, dass diese Modifikation weiterhin korrekt die SZKs ermittelt, und dies in derselben Reihenfolge, in der es der originale Tarjan-Algorithmus tut. Da die Exploration des Graphens und die Erkennung der abgeschlossenen SZKs über die Lowlink-Nummern unberührt bleibt, genügt es zu zeigen, dass zu diesen Zeitpunkten tatsächlich weiterhin alle Elemente der SZK oben auf dem Tarjan-Stack liegen und auch genau diese abgeräumt werden.

Im Folgenden wird mit  $[a]$  die SZK des Knotens  $a$  bezeichnet. Weiterhin definieren wir eine Halbordnung  $\leq$  auf den SZKs, sodass  $A \leq B$  gilt gdw. es einen Pfad von SZK  $A$  nach SZK  $B$  gibt.

**Observation 1** *Die DFS-Nummern jedes Knoten einer SZK ist mindestens so groß wie die DFS-Nummer des Knotens, mit dem die SZK betreten wurde.*

**Lemma 2** *Wenn auf dem modifizierten Tarjan-Stack Knoten  $a$  unterhalb von Knoten  $b$  liegt, gilt  $[a] \leq [b]$ .*

*Beweis.* Angenommen es gäbe einen Knoten  $a$  unterhalb von  $b$  mit  $[a] \not\leq [b]$ . Sei  $a^*$  derjenige Knoten, mit dem die SZK  $[a]$  initial betreten wurde. Nachdem  $a^*$  betreten wurde, kann  $b$  nach Annahme nicht exploriert oder abgeschlossen werden, bevor  $a^*$  verlassen wird. Alle Knoten, die von  $a^*$  aus exploriert werden, also insbesondere  $a$ , haben eine mindestens ebenso große DFS-Nummer wie  $a^*$ . Nach dem Verlassen von  $a^*$ , also nach dem Abschluss von  $a$  aber vor dem Abschluss von  $b$ , wurde die SZK  $[a]$  vollständig exploriert und alle Elemente mit einer mindestens ebenso großen DFS-Nummer wie  $a^*$  werden vom Stack entfernt, insbesondere  $a$ . Dies ist ein Widerspruch zur Annahme, dass sich  $a$  noch auf dem Stack befindet, wenn  $b$  abgeschlossen wird.  $\square$

**Corollary 3** *Auf dem modifizierten Tarjan-Stack liegen alle Knoten, die zu derselben SZK gehören, zu jedem Zeitpunkt zusammenhängend auf dem Stack.*

**Theorem 1** *Für jede gefundene SZK werden genau die Elemente der SZK vom Tarjan-Stack entfernt.*

*Beweis.* Immer wenn eine SZK  $A$  abgeschlossen wird, ist das oberste Element auf dem modifizierten Tarjan-Stack derjenige Knoten  $a^*$ , mit dem die SZK initial betreten wurde. Nach Korollar 3 liegen also alle Knoten von  $A$  zusammenhängend oben auf dem modifizierten Tarjan-Stack. Außerdem gilt nach Lemma 2 für jeden Knoten  $x$  auf dem Stack  $[x] \leq A$ , und folglich besitzt jeder Knoten auf dem Stack, der nicht in  $A$  liegt, eine kleinere DFS-Nummer als  $a^*$ . Alle Knoten in  $A$  besitzen hingegen nach 1 eine mindestens ebenso große DFS-Nummer wie  $a^*$ , es werden also exakt die Elemente der SZK  $A$  vom modifizierten Tarjan-Stack abgeräumt.  $\square$