Language theoretic comparison of LL and LR grammars

Asked 8 years, 10 months ago Active 3 years, 8 months ago Viewed 12k times

69

38

People often say that LR(k) parsers are more powerful than LL(k) parsers. These statements are vague most of the time; in particular, should we compare the classes for a fixed k or the union over all k? So how is the situation really? In particular, I am interested in how LL(*) fits in.

As far as I know, the respective sets of grammars LL and LR parsers accept are orthogonal, so let us talk about the languages generated by the respective sets of grammars. Let $LR(k)$ denote the class of languages generated by grammars that can be parsed by an $LR(k)$ parser, and similar for other classes.

I am interested in the following relations:

- $LL(k) \overset{?}{\subseteq} LR(k)$
- $\bigcup_{i=1}^{\infty} LL(k) \overset{?}{\subseteq} \bigcup_{i=1}^{\infty} LR(k)$
- $\bigcup_{i=1}^{\infty} LL(k) \overset{?}{=} LL(*)$
- $LL(*) \overset{?}{\circ} \bigcup_{i=1}^{\infty} LR(k)$

Some of these are probably easy; my goal is to collect a "complete" comparison. References are appreciated.

edited Jun 6 '13 at 14:05
frafl
2,240 1 12 32

asked Mar 7 '12 at 0:32
Raphael ♦
68.9k 27 156 350

2 Maybe this can help you! Grammar Hierarchy image – Andrea Tucci Jul 11 '12 at 15:32

1 @AndreaTucci: Yes, but that only covers the grammars, not the generated languages. – Raphael ♦ Jul 11 '12 at 18:06

1 Answer

Active Oldest Votes

62

There are numerous containments known. Let $\subseteq$ denote containment and $\subset$ proper containment. Let $\times$ denote incomparability.
Let $LL = \bigcup_k LL(k)$, $LR = \bigcup_k LR(k)$.

Grammar level

For LL

- $LL(0) \subset LL(1) \subset LL(2) \subset LL(2) \subset \cdots \subset LL(k) \subset \cdots \subset LL \subset LL(*)$
- $SLL(1) = LL(1), SLL(k) \subset LL(k), SLL(k+1) \times LL(k)$

Most of these are proven in Properties of deterministic top down grammars by Rosenkrantz and Stearns. $SLL(k+1) \times LL(k)$ is a rather trivial exercise. This presentation by Terence Parr places $LL(*)$ on slide 13. The paper LL-regular grammars by Jarzabek and Krawczyk show $LL \subset LLR$, and their proof trivially extends to $LL \subset LL(*)$

For LR

- $LR(0) \subset SLR(1) \subset LALR(1) \subset LR(1)$
- $SLR(k) \subset LALR(k) \subset LR(k)$
- $SLR(1) \subset SLR(2) \subset \cdots \subset SLR(k)$
- $LALR(1) \subset LALR(2) \subset \cdots \subset LALR(k)$
- $LR(0) \subset LR(1) \subset LR(2) \subset \cdots \subset LR(k) \subset \cdots \subset LR$

LL versus LR

- $LL(k) \subset LR(k)$ (Properties of deterministic top down grammars plus any left recursive grammar)
- $LL(k) \times SLR(k), LALR(k), LR(k-1)$ (simple exercise)
- $LL \subset LR$ (any left recursive grammar)
- $LL(*) \times LR$ (left recursion versus arbitrary lookahead)

Language level

For LL

- $LL(0) \subset LL(1) \subset LL(2) \subset \cdots \subset LL(k) \subset \cdots \subset LL \subset LL(*)$
- $SLL(k) = LL(k)$

Most of these are proven in Properties of deterministic top down grammars. The equivalence problem for LL- and LR-regular grammars by Nijholt makes references to papers showing $LL(k) \subset LL(*)$. The paper LL-regular grammars by Jarzabek and Krawczyk show $LL \subset LLR$, and their proof trivially extends to $LL \subset LL(*)$

For LR

- $LR(0) \subset SLR(1) = LALR(1) = LR(1) = SLR(k) = LALR(k) = LR(k) = LR$

Some of these were proven by Knuth in his paper On the Translation of Languages from Left to Right in which he introduced LR(k), the rest are proven in Transforming LR(k) Grammars to LR(1), SLR(1), and (1,1) Bounded Right-Context Grammars by Mickunas et al.

LL versus LR

- $LL \subset LR(1)$ (containment follows from the above, $\{a^i b^j | i \geq j\}$ is the canonical example for strict containment)
- $LL(*) \times LR$ (the language $\{a^i b^j | i \geq j\}$ shows half the claim, and the introduction of The equivalence problem for LL- and LR-regular grammars by Nijholt makes references to papers showing the other half)
- $LR(1) = DCFL$ (see e.g. reference here).

1 8,676 2 31 62

Excellent answer though, I had already upvoted. I would have thought Frank deRemer proved LALR <= LR in his original LALR paper? (1969?) – user207421 Apr 11 '12 at 18:43

@EJP: $LALR(k) \subseteq LR(k)$ is immediate if you define $LALR$ by collapsing LR states. deRemer only proved that his construction created the same parser. The grammar classes are not equal, which is a nice exercise (or even a nice question for this site). The language class equality is a lot more tricky: read the paper if you really want the details ;) – Alex ten Brink Apr 11 '12 at 22:58

1 @AlextenBrink I did read the paper, and was taught by Frank de Remer, but it's 30+ years ago ;-) Thanks for all the detail. – user207421 Apr 12 '12 at 0:13

It might be nice to collect example grammars for each inequality. – o11c May 31 '15 at 5:51

1 @o11c I think that would overburden a single answer. My impression is that Alex gave good references where necessary; he states "easy exercise" for some. I guess if a reader can not come up with a grammar, they can post a new question asking for that specific case. – Raphael ♦ Jun 18 '15 at 10:33