Hogyan szerkesszünk automatát bonyolult feladat esetén?

A feladat

Az L nyelv alfabetája $\Sigma = \{\mathtt{a}, \mathtt{b}\}$. A nyelv mondatai azok a nem üres jelsorozatok, amelyekben van legalább egy páros hosszúságú teljes homogén részsorozat. Szerkessz minimálautomatát a következő nyelvekre:

$$L \quad L^2 \quad L \cap L^2 \quad L^*$$

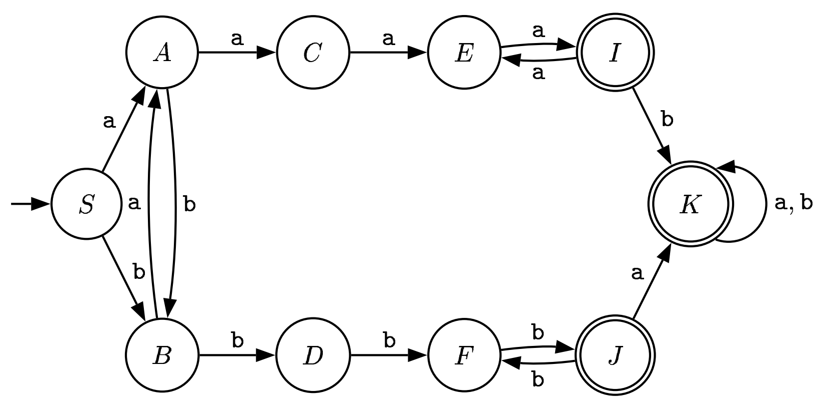
A megoldás

A megoldás most nem lesz teljes körű, mivel csak az L^2 és az $L \cap L^2$ nyelv automatája kerül megszerkesztésre. Nézzük a különböző eseteket!

- Ha a jelsorozatban felváltva jönnek az `a` és `b` karakterek, akkor semmivel sem jutunk közelebb a célunk felé. Vagy másképpen úgy is fogalmazhatnánk, hogy az A és a B állapotban még nem volt legalább kettő hosszú teljes homogén részsorozat. A továbbiakban az esetek tárgyalásánál mindig azt fogom feltételezni, hogy az első, legalább kettő hosszú teljes homogén részsorozat a mondat legelején található. (Ezt természetesen úgy kell értelmezni, hogy lehet a szó elején `a`-kból és `b`-kből álló részsorozat, amennyiben ezek felváltva szerepelnek. Ez az automatában az A és a B állapotok közti mozgásnak felel meg.)

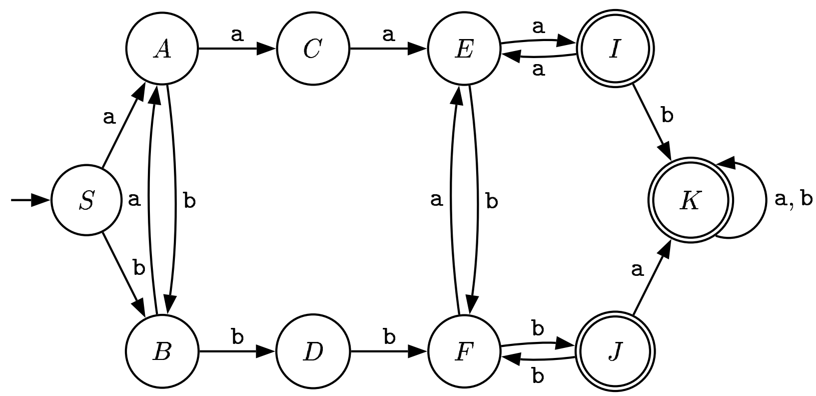
  A feladat szimmetrikusságát már rögtön érezhetjük, mivel semmilyen, betűkre vonatkozó kritériumunk nincsen. Tehát a továbbiakban az általánosság csorbítása nélkül feltételezhetjük, hogy a jelsorozat `a`-val kezdődik.

- Ha az adott jelsorozat $\mathtt{aaaa}(\mathtt{aa})^*$ alakú, akkor azt értelemszerűen el kell fogadnunk, mivel $\mathtt{aa}|\mathtt{aa}(\mathtt{aa})^*$ módon a szó szétvágható két részre úgy, hogy mindkét részszó L-beli. Ha a fenti módon leírt jelsorozatok végére még teszünk egy `b`-t, és utána tetszőleges módon folytatjuk, akkor az így kapott szavakat is elfogadjuk, mivel az első része már megfelel a kívánalmaknak.



  Az eddig elmondottak tömör megfogalmazása valahogy úgy hangzik, hogy a jelsorozatban van egy legalább négy hosszú, páros, teljes homogén részsorozat.

- Most akkor nézzük az $\texttt{aaa(aa)}^*$ szituációt! (A jelsorozat eddigi részében van egy legalább három hosszú, páratlan teljes homogén részsorozat.) Ez szétvágható $\texttt{aa}|\texttt{a(aa)}^*$ módon. Az első részszóba ekkor kerül két darab a, míg a másik részszó továbbra is páros, teljes homogén részsorozat nélkül árválkodik (pl. aa|aaab). Ezután már csak ennek a megérkezésére kell várnunk. (Vegyük észre, hogy az automatánk E, F, I, J, K állapotokból álló része feltűnően hasonlít az L nyelvet elfogadó automatára (vagy pontosabban annak „hátsó" részére), amikor szintén a páros, teljes homogén részsorozatra vártunk!) Tehát most kb. itt tartunk:



- Most már csak a legkülönösebb eset maradt hátra (C állapot). Eddig jött pontosan két darab a karakter, majd most jön egy b. Úgy tűnik, hogy ilyen esettel eddig még nem találkoztunk. Praktikus ezért bevezetni egy új állapotot (G).

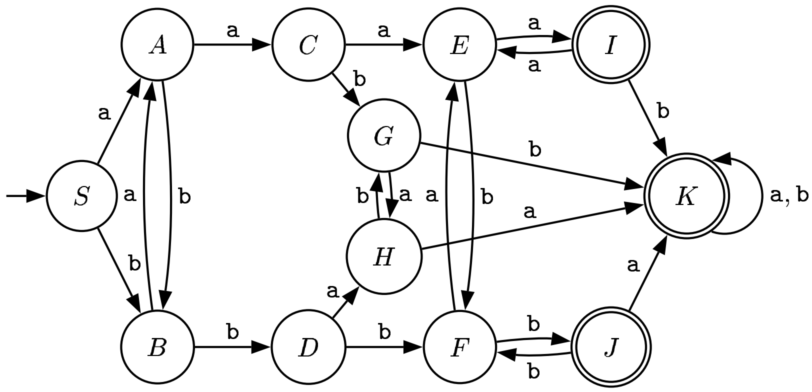
  - Ha a két a után páros sok b jön, akkor a szót el kell fogadnunk, mivel a szó $\texttt{aa}|\texttt{bb(bb)}^*$ alakban szétbontható két részszóra úgy, hogy mindkét részszó L-beli. Természetes akkor is elfogadás a válaszunk, ha a páros sok b-t tetszőleges karaktersorozat követi ($\texttt{aa}|\texttt{bb(bb)}^*\texttt{a(a} + \texttt{b)}^*$).
  - Ha a két a után páratlan sok (és legalább három) b jön, akkor a szót ismét elfogadjuk, mivel a feldarabolás $\texttt{aab}|\texttt{bb(bb)}^*$ módon megoldható. Ha az ilyen alakú szavakat még folytatjuk, az elfogadás ténye nem fog megváltozni.

  Az előbbi két pontban foglaltakat összefoglalhatjuk úgy, hogy ha a kezdeti két a után jön még legalább két b, akkor elfogadó állapotba kell mennünk, és ott is kell maradnunk, akármilyen karakter is jön ezután. Összefoglalhatjuk ezt így is:

$$\delta(G, \texttt{b}) = K$$

  Amennyiben aab után a jön, akkor G-ből a szimmetrikus párjába (H) kell menni. (Pl. az aabaa szót megnézve ez természetesnek tűnik.)

Minden eddigit összefoglalva automatánk a következőképpen néz ki:



Tehát az L^2 nyelv automatájával készen vagyunk. (Pontosabban még kéne ejteni néhány szót arról, hogy az automata minimális, de ezt most kihagyom.)

Az $L \cap L^2$ nyelv automatájának szerkesztését gyorsabban meg tudjuk oldani.

1. Egyrészt $L \cap L^2 \subseteq L^2$ triviálisan igaz (L^2 mondataira még további megkötéseket teszünk).

2. Másrészt pedig azt kell észrevenni, hogy ha $w \in L^2$, akkor $w \in L$, azaz minden L^2-beli mondat egyben L-beli is. Vagy másképpen írhatjuk így is: $L \supseteq L^2$, illetve

$$L \cap L^2 \supseteq L^2 \cap L^2 = L^2$$

   Ennek igazolását a szorgos Olvasóra bízom.

Tehát $L \cap L^2 = L^2$. Ebből következően a két nyelv minimálautomatája ugyanaz.