

Âúââèæáíèà è ìðíâðèà àèìòàç â ñèñòàíà ðàñííçíàâàíèý äðâáíáðóññèíè ñèìðíèñè

Ààðìð Èàáí Àíàðìèùààè- Çàèáíòíà
10.07.2010 ä.
İñèàáíáà íáííàèáíèà 10.07.2010 ä.

Electronic processing of ancient Russian text requires its representation in electronic text format. An OCR system may help in translation of real documents into electronic text. A structure-based hypothesis-driven recognition method is proposed. The paper describes the principles of hypotheses nomination and testing used by this method.

Äèý ððàíáíèý, ýèàèððííííàí èççàáíèý è èíííðòàððííè íáðâáíðèè ìàìýòíèíà äðâáíáðóññèíè ñèìðíèñè ððâáóòòíý èð ìðââñòàèèá çíáíèè. Ýòí ìáñòíýòàèèñòàí ààèààð òàèàñííàðçàçííí ñíçàáíèà ààðííàðèçèððíáííè èíííðòàððííè ñèñòàíú ðàñííçíàâàíèý àíèòíáíòíà ýò ðàñííçíàâàíèý, ìñòðíáííàý íà ñèàáóðòèð ðàðâèðððííó ìðèèèèàð: ñòðòèèòòðííè ìððíà è ðàñííçíàâàíèèð [Ñòðòèèòòðííà 1977], ððàñí àáóòòðíáííàâý ìðâáíèçàèèè ìðíòàññà ðàñííçíàâàíèý (ñèíáà è áóèèú), íá-, òèíà ìíèñàíèà ñòðòèèòòðííó ýèàíáíòíà ðàñííçíàâàííó ìðââèèàáàòòíý èñèèð-àòú ìóð, ì òíðàèèáíèý ìðíòàññíí ðàñííçíàâàíèý àúâàèæáíèàí àèìòàç í ðàñííçíàâàííó ìáúáèòàð è èð ìðíáðð ìáðçíí, àúâàèýòòíý òíèèè ñóúáñòàáííúà -àñòè èçíáðâæáíèý, à ìñòíðííèèà ìñòàòòòíý áàç áíèàíèý. Ðàçðâáíòàí ñíííá ìíèñàíèý ñ 1979]. Ñòðòèèòòðà ñèíà è áóèèà èàðâðòè-àñèè ìáðçíí ìðââñòàèèýàòòíý à áàçà çíáíèè ñèñòàíú. Ìñíáííà óçèú ððàèíáíè ñàòè èí ìáðâá-áíèý. Ðàñííçíàâàíèà áóèèà è ñèíà ñòðíèòòíý íà ìñíáà èííòàíòèè àèðòàèèííàí òðàèíà. Äèý ìðââèè, ìííòè ðàñííòòèè ìðíòàâ ìíèñòàòòàèè ìáàèðáàáíòð èàðòèíó. Ýòíò ððàèí ìàçúâàáòòíý àèðòàèèííà. Çàâà-à ðàñííçíàâàíèý ñáíàèòòíý è ìáðíæááíèè ñíííáà ñòðòèèòòðà àèðòàèèííàí òðàèíà ìçáíèýàð àúâàèèòú ìááíð òðàèíá áàçú çíáíèè à èà-àñòàà ñíèñèà àèìòàç, ìðáíòèèèíí ìíèñòàâ ññòèíè íà ñíæàñíáííúà óçèú è òáí ñáíúí ìíèñòàòòàý ñòáíó ñíæàñíáíèý. Äèý ìðââèèáíèý ìðââñííáííòè àèìòàç èñííèèçòòòò ìðââñèèèèèè òðàèíó è ìðââèèýàòòíý èàè ìòííáíèà -èñèà ìáð ñíæàñíáííúò óçèíà à àèìòàçà è ìáúáíò -èñèó óçèíà à òðàèíà á àèìòàçú àíáíðèò í òí-ìíòè òñòàííàèáííàí ñíòàòòòàèè è àú-èñèýòòíý èàè ìòííáíèà -èñèà ìáð à àèìòàçà è -èñèó óçèíà ÁÓ. àèìòàçú. Ìðè ìðíáððèà àèìòàç àèý èàæáíè èç íèð àú-èñèýòòíý ñòáíáíè ñíæàñíáíííòè è ìèèáíííòè è ìáíà èç íèð ìðèíèààòò èèíèè. Ààèàà ððàñíèððíáíèèòò àðíáííàí èçíáðâæáíèý ìðââèèòòíý çáíðííú íà ìíèñè ìæèàááíúò èèíèè íà èçíáðâæáíèè, ìíèó-áííúà ðð òðàèíàíè áàçú çíáíèè àí àñàò àèìòàçàð. Ààèàà àúñíèýàòòíý ìáðâá-, ò ðàðâèðððèñòèè àèìòàç è ìáðòòèèèèà òñèíáèà ìðèáíííò òèàçúâàáíàý àáííè àèìòàçíè. Áñèè òàèíè àèìòàçú íáò, òàèóúáè íàçíà-ààòòíý àèìòàçà ñ ìàèñèíàèèííè ñòáíáííð ñíæàñíáíííòè ñíæàñíáíèà óçèíà áóèèà è ìðíòòàííòàáííúò ìòííáíèè ìæáò íèèè. Èèòàðâòòòà1. xáðáííèí È.Á. Ðóññèàý Ìàèáíáðàòèè Ì., 1956 ðàñòðíáúò ìííòòíííúò èçíáðâæáíèè àèàðèáííàí ðáààèòíðà SpotLight // Èíòàèèèèèèèèèèèèèèèèèè è ñèñòàíú: ñáíðíèè ñòàòàè àèáíðèòíú àíàèèçà èííòòííúò èçíáðâæáíèè à àèçóàèèííúò èííòíàòèèííúò ñèñòàíàð íà ìñíáà ìáíáííòíáííè íá-, òèíè ñáíáíòè-àñèíè ñá 2004.5. Ìèíèè Ì. Òðàèíú àèý ìðââñòàèèáíèý çíáíèè. Ìáð. ñ áíàè. 1979.