

A duplakocka modell és az igei szerkezeteket kinyerő „ugrik és marad” módszer

PPKE ITK

Sass Bálint

Nyelvtudományi Kutatóközpont

`sass.balint@nytud.hu`

motiváció

A nyelv alapegységei nem a szavak, hanem a szerkezetek.
A szó csak szélső eset: olyan szerkezet, ami egy elemből áll.
:)

Simán lehet, hogy adott szóra valamely nyelven
csak szerkezet van!

magyar *'egymás'* ↔ angol *'each other'*
wolof *'gëmm'* ↔ magyar *'behunyja a szemét'*
magyar *'krumpli'* ↔ francia *'pomme de terre'*

szótár → szerkezettár

alapok = modell & módszer

példa:

‘Részt vesz az akcióban.’

vesz ACC : rész INE : akció

alapok = modell & módszer

példa:

‘Részt vesz az akcióban.’

vesz ACC : rész INE : akció

→ ige...

alapok = modell & módszer

példa:

‘Részt vesz az akcióban.’

vesz **ACC** : rész **INE** : akció

→ ige + helyek...

alapok = modell & módszer

példa:

‘Részt vesz az akcióban.’

vesz ACC : **rész** INE : **akció**

→ ige + helyek + kitöltők

alapok = modell & módszer

példa:

‘Részt vesz az akcióban.’

vesz ACC : rész INE : akció

→ ige + helyek + kitöltők

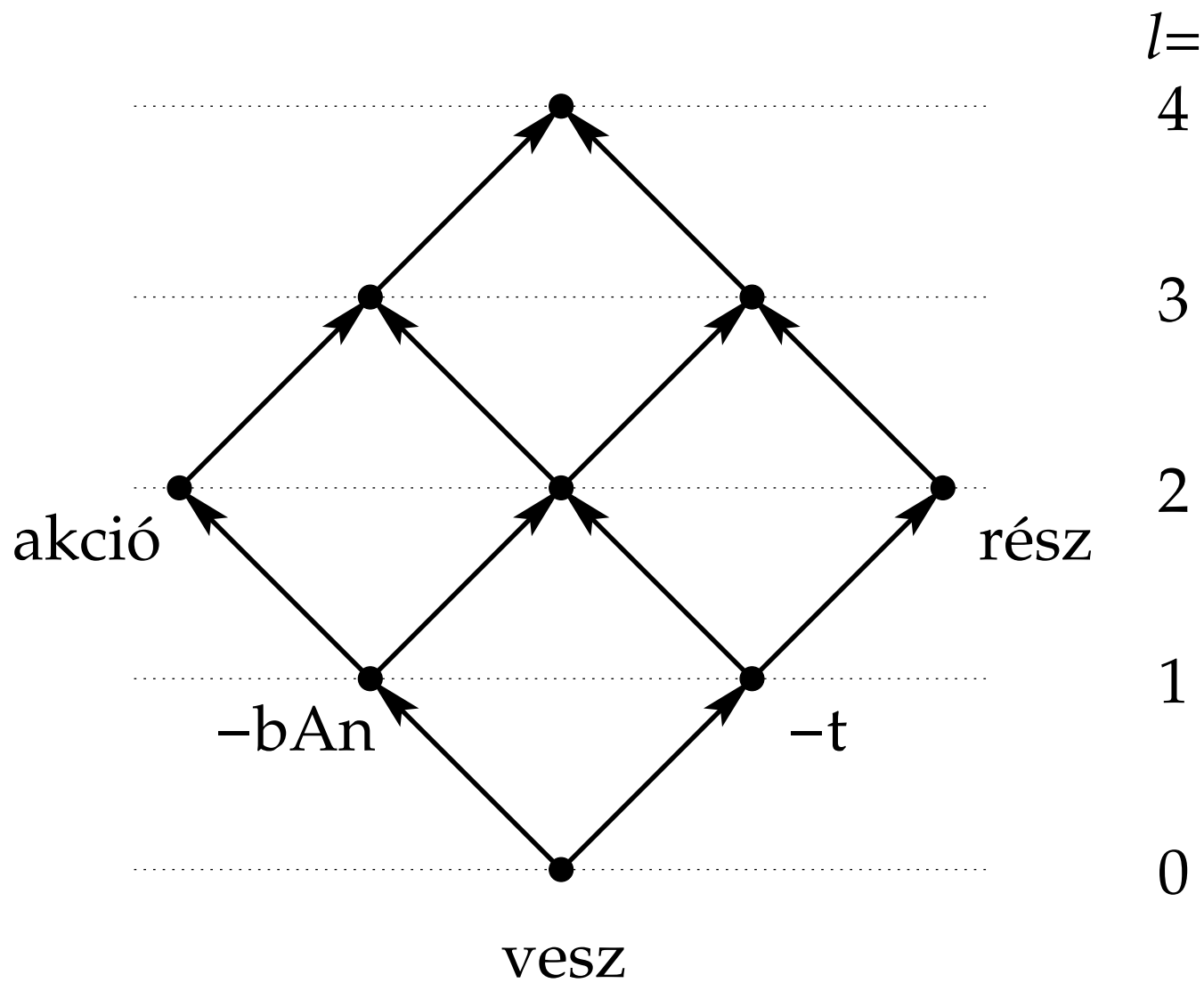
cél:

‘vesz részt vmiben’

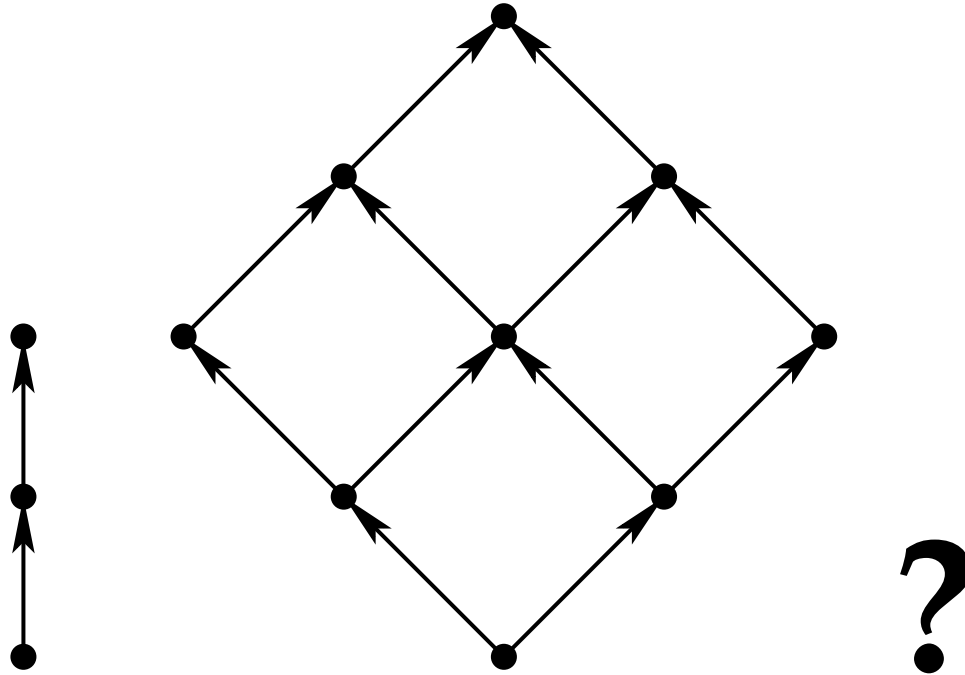
= teljes és tiszta **valódi igei szerkezetek**

= lexikográfiailag hasznos igei szerkezetek

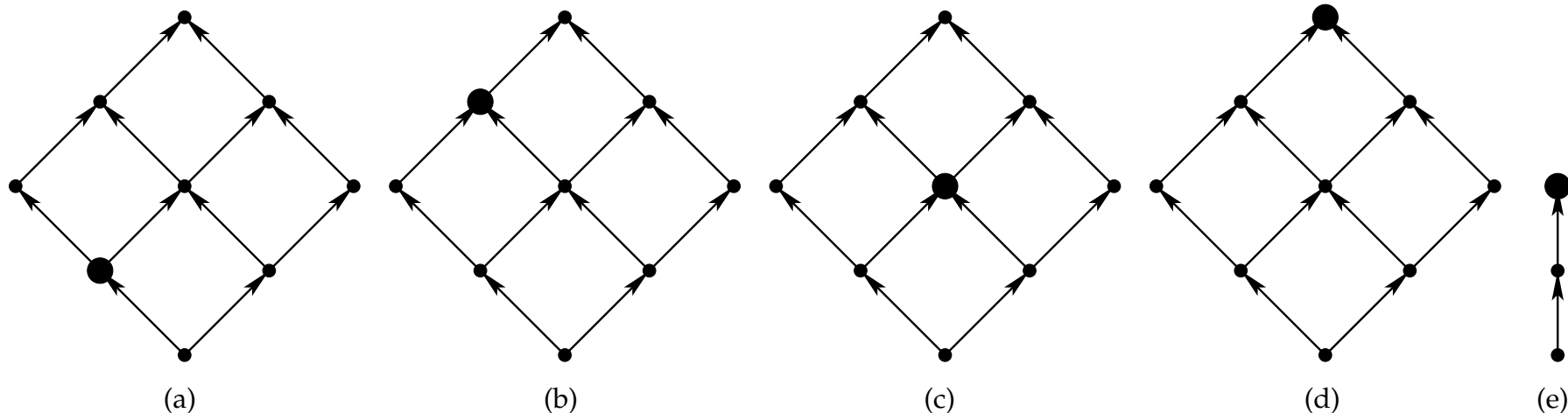
duplakocka



bővítmények = dimenziók



valódi igei szerkezet



(a) '*olvas könyvet ágyban*' → '*olvas vmit*' (alany \emptyset)

(b) '*vesz részt felolvasáson*' → '*vesz részt vmin*' (alany \emptyset)

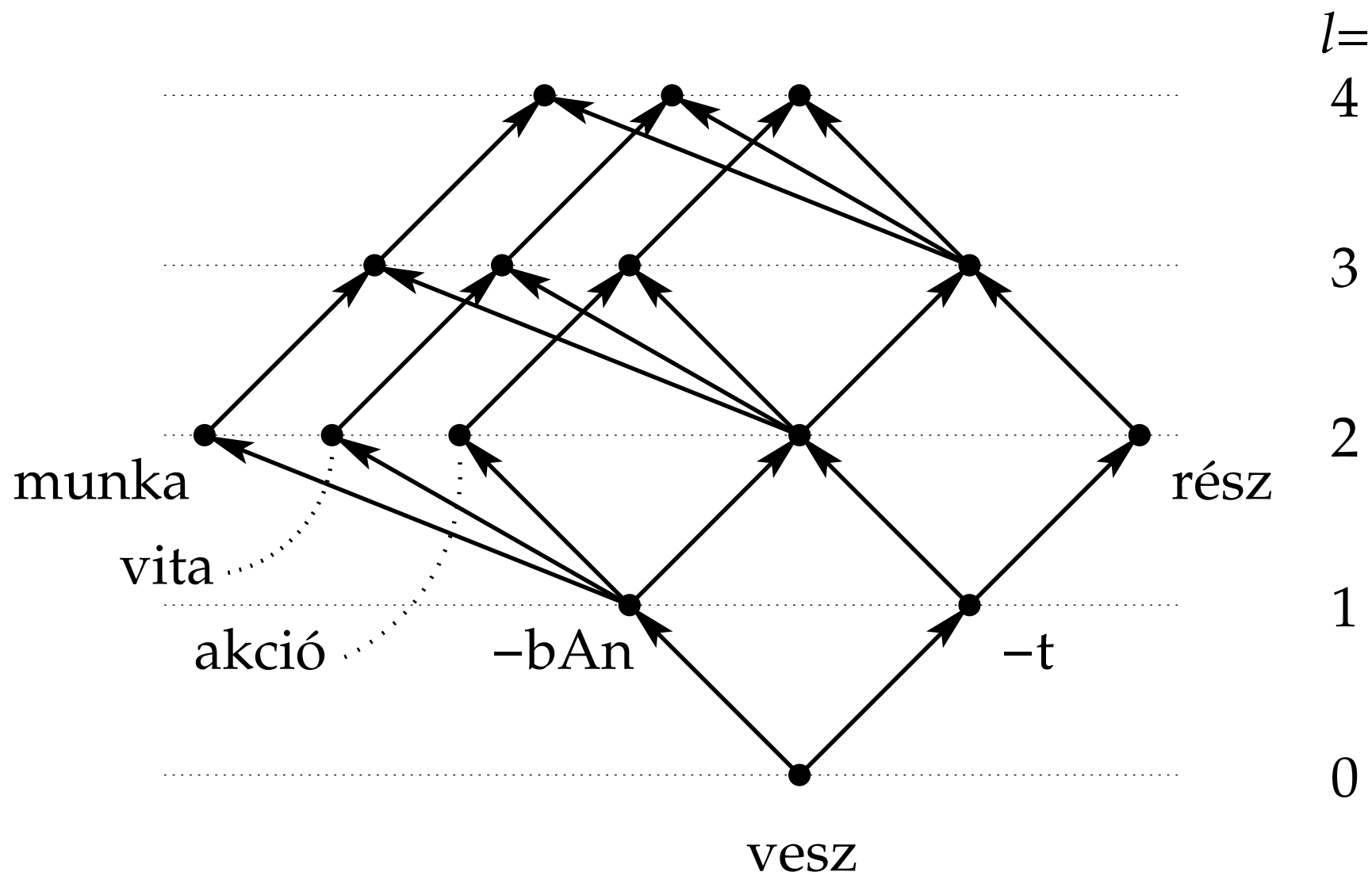
'*kerül sor megbeszélésre*' → '*kerül sor vmire*'

(c) '*ad Jánosnak könyvet*' → '*ad vkinek vmit*' (alany \emptyset)

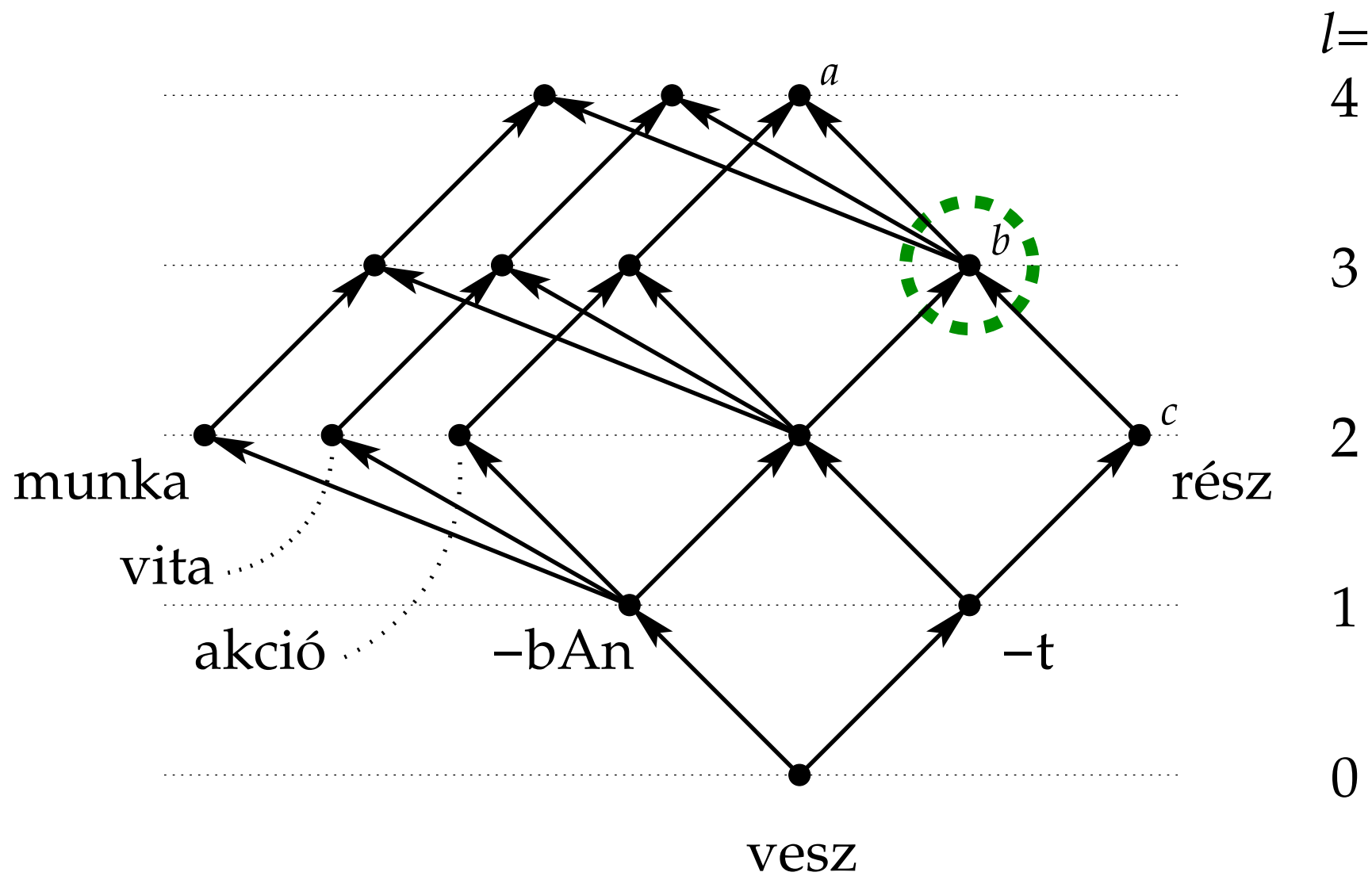
(d) '*Ne fessd az ördögöt a falra!*' → '*fest ördögöt falra*' (alany \emptyset)

(e) '*Kisütött a nap.*' → '*kisüt nap*'

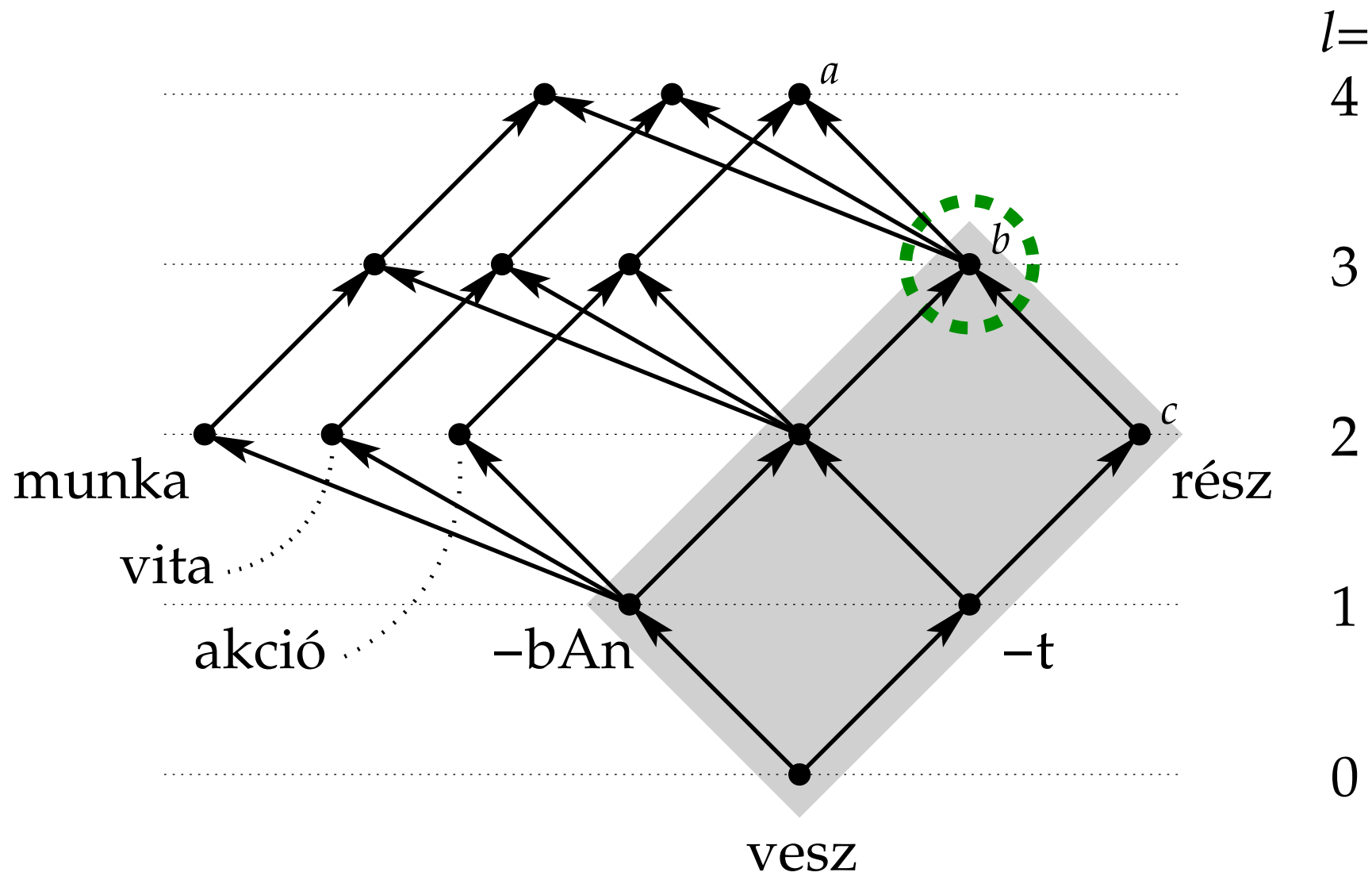
korpuszháló



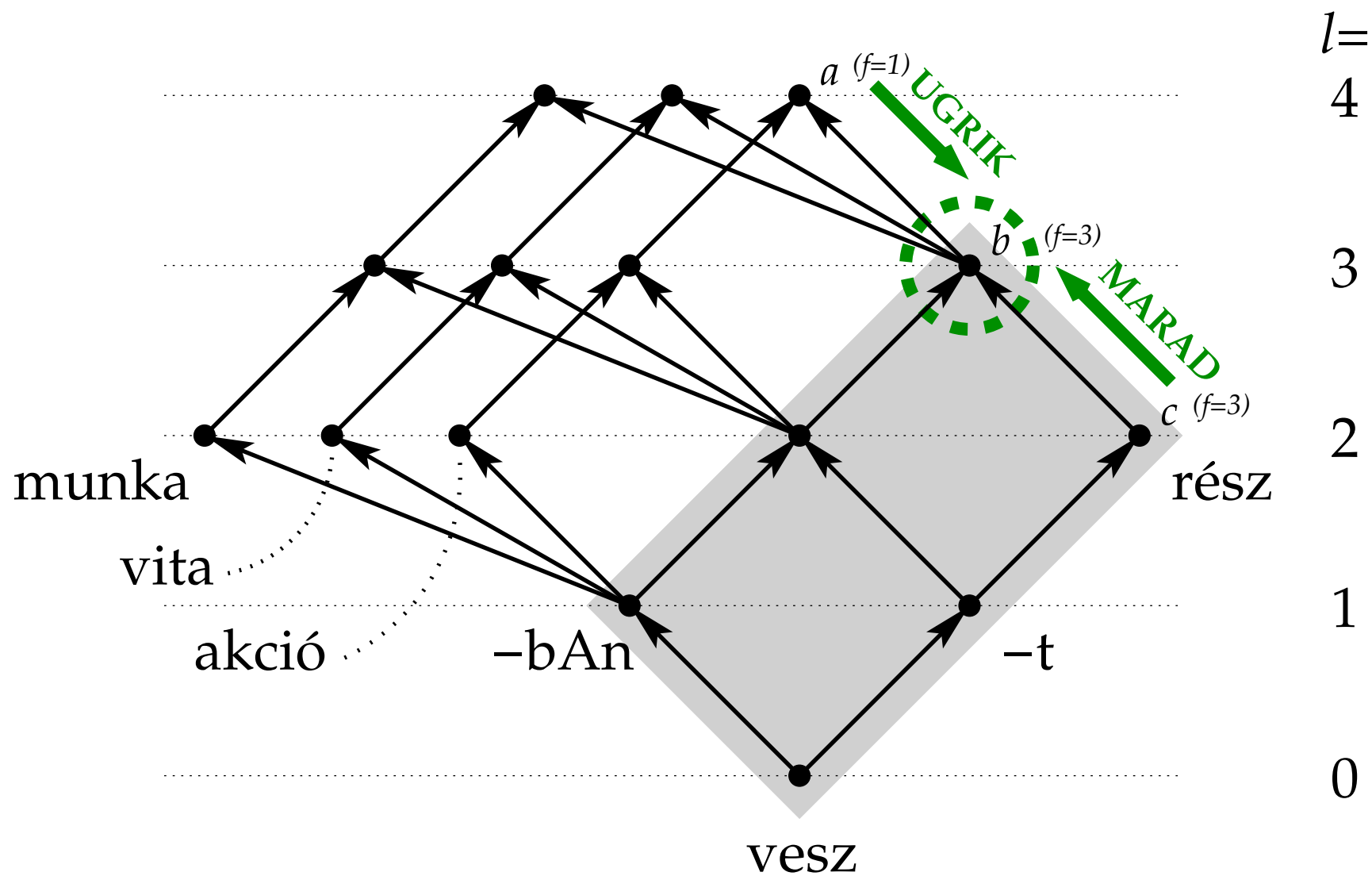
valódi igei szerkezet

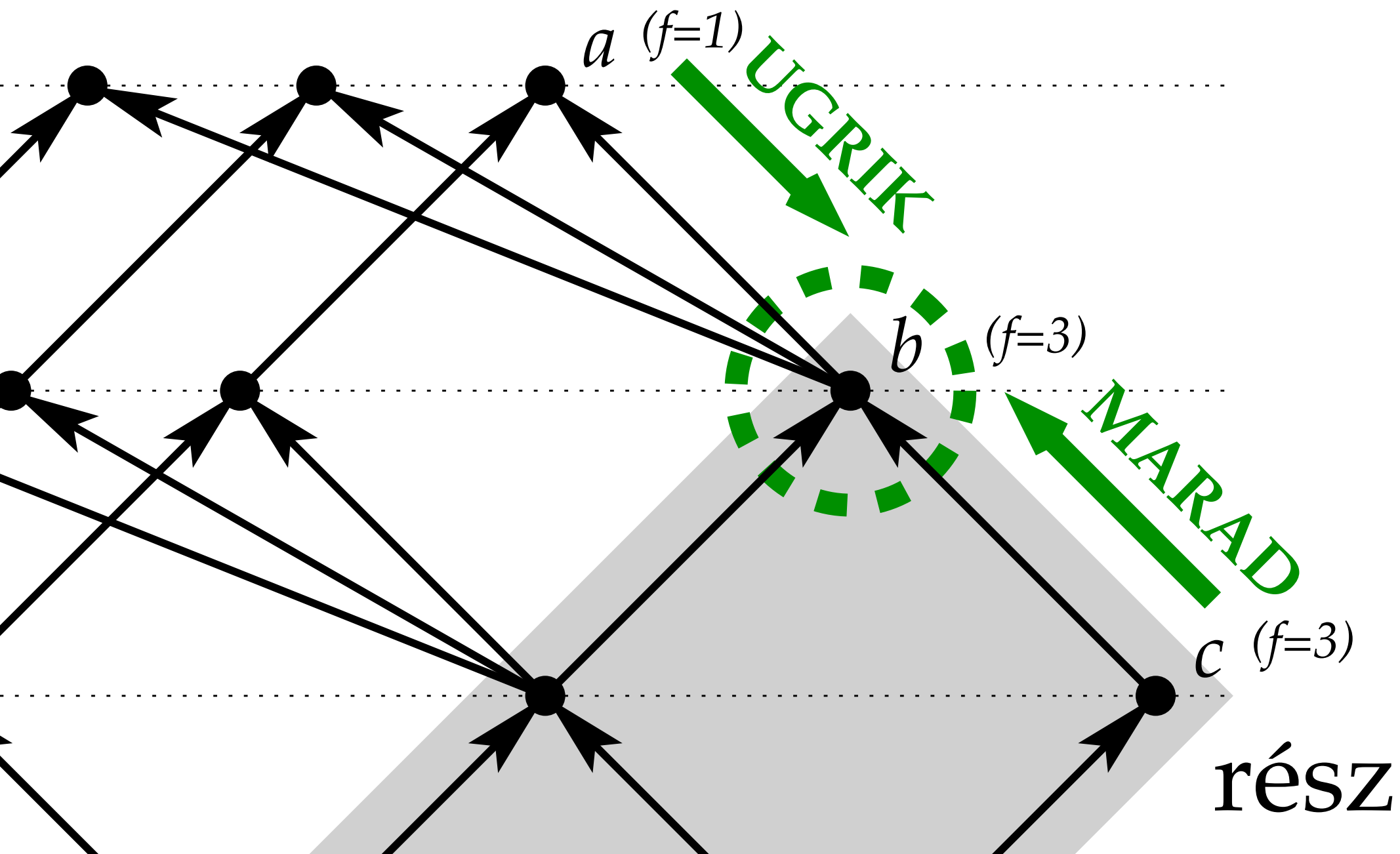


f függvény



„ugrik és marad” módszer





aktuális kutatási kérdések

- (1) a módszer tetszőleges nyelvre alkalmazható? *Igen.*
- (2) a módszer alkalmazásához szükséges adatok
függőségileg elemzett korpuszból könnyen származtathatók? *Igen.*

nyersanyag: UD korpuszok

cseh, német, angol, finn, magyar, holland, norvég, török, wolof

eredmények

#	nyelv	igei szerkezet	magyar megfelelő
1.	cs	být SUBJ:rozdíl mezi	(van különbség vmi között)
2.	cs	investovat do	(befektet vmibe)
3.	cs	čekat se	(várandós)
4.	de	fallen SUBJ:aktie auf	(esik részvény vmire)
5.	de	finden sich SUBJ:info auf	(megtalálható információ vhol)
6.	de	handeln sich um	(arról van szó)
7.	en	do IOBJ OBJ:favor	(szívességet tesz vkinek)
8.	en	get in:touch with	(kapcsolatba lép vkivel)
9.	en	make sure	(meggyőződik)
10.	en	take OBJ:care of	(vigyáz vmkire)
11.	fi	ottaa Ill :huomio OBJ	(figyelembe vesz vmit)
12.	fi	ottaa Ill :käyttö OBJ	(használatba vesz vmit)

eredmények

#	nyelv	igei szerkezet	magyar megfelelő
13.	fi	ottaa Ill:käsi OBJ	(kézbe vesz vmit)
14.	hu	lesz SUBJ:szükség -rA	
15.	hu	tesz lehetővé -t	
16.	nl	zien OBJ:kans te	(lát lehetőséget vmit csinálni)
17.	no	få OBJ med:seg	(magával visz vmit)
18.	no	få OBJ:gjennomslag i	(áttörést ér el vmiben)
19.	no	få OBJ på:seg	(felvesz vmit (ruhafélét) magára)
20.	no	få OBJ:tillit	(visszanyeri az önbizalmát)
21.	no	få OBJ i:løp	(futtat vmit (szoftvert))
22.	no	ha OBJ på:seg	(vmi (ruhaféle) van rajta)
23.	wo	am OBJ:kàttan ci	(van energiája vmit csinálni)
24.	wo	wax IOBJ OBJ	(mond vkinek vmit)

eredmények

mind jó:

look → look, look good / great, look like

deal → deal with

go → go to (fn), going to (ige), go crazy

szerkezetek kontrasztív megfeleltetése:

cs mluvit/hovořit o

de sprechen von

en talk about

fi puhua Ela (-stA)

hu beszél Del (-rÓl)

nl praten over

no snakke om

eredmények

helyek (slot-ok) kontrasztív megfeleltetése:

cs	de	magyar megfelelő
čekat Acc=na Acc=na	warten Acc=auf Acc=auf	(vár -rA)
mít Dat=k:dispozice patřit Dat=k Dat=k	stehen Dat=zu:Verfügung gehören Dat=zu Dat=zu	(rendelkezésre áll) (tartozik vmihez)

irányok, ötletek, tervek

1. jobb módszer, hibák kiküszöbölése: *'få øye på løve'* ↔ *'húz ACC közé'*
2. nagy (nem magyar) dep elemzett korpuszra futtatni: DEPCC
3. plusz egy szint: *'fontos szerepet játszik vmi vmiben'*
4. minden root kezelése → névszói állítmányok
5. kiértékelés ... de hogyan?
6. triplakocka – pl. élő/élettelen
7. klaszterezés – *'eszik ACC:<ÉTEL>'*
8. igeiszerkezet-embedding – *'participate in' ~ 'take part in'*
9. hely-embedding (slot-embedding):
'få OBJ på:seg' ~ 'ha OBJ på:seg', 'fő NOM' ~ 'főz NOM ACC'
- ...

összefoglalás

igei szerkezetek

valódi igei szerkezetek

modell: duplakocka

kinyerő módszer: „ugrik és marad”

nyelvfüggetlen (predikátum-argumentum struktúra kell hozzá)

→ lexikográfiailag is hasznos valódi igei szerkezeteket kapunk

elérhető:

`github.com/sassbalint/double-cube-jump-and-stay-multilingual`

olvasnivaló: Sass Bálint. A duplakocka modell... In: *MSZNY 2020*.

Sass Bálint – `sass.balint@nytud.hu`