

Computational Linguistics 2 (CL3.202)

Monsoon 2021, IIIT Hyderabad

Assignment 1

Question 1

Semantic role labelling is the annotation of a text with the semantic roles (as labels) of the referring expressions in the text.

Semantic roles (or *thematic roles*) indicate the relation that a referring expression holds to the rest of the sentence. The most common set of roles includes:

- agent: the performer of the action
- experiencer: the entity feeling/experiencing something
- force: the involuntary cause of the occurrence
- theme: the entity most affected by the occurrence
- result: the final product of the process
- content: the proposition involved in the action
- instrument: the entity used to perform the action
- beneficiary: the entity for whose benefit the action was performed; the recipient
- source: the place of origin of a moving entity
- goal: the destination of a moving entity

Examples of semantic role labelling:

Ravi was feeling sleepy
EXPERIENCER PREDICATE

Ram gave Shyam a book
AGENT PREDICATE BENEFICIARY THEME

The water flowed from the dam to the ocean
AGENT PREDICATE SOURCE GOAL

The rain wet the roof
FORCE PREDICATE THEME

They asked him a question
AGENT PREDICATE BENEFICIARY CONTENT

Question 2

The following five sentences have been syntactically analysed and semantic role labelling has been done on them.

1. **She went to school by car.**

Syntactic analysis:

[_S [_{NP} She] [_{VP} [_V went] [_{PP} [_P to] [_N school]] [_{PP} [_P by] [_N car]]]]

Semantic analysis:

She went to school by car
AGENT PREDICATE GOAL INSTRUMENT

2. **The storm broke the door.**

Syntactic analysis:

[_S [_{NP} [_{Det} The] [_{NP} storm]] [_{VP} [_V broke] [_{NP} [_{Det} the] [_N door]]]]

Semantic analysis:

The storm broke the door
FORCE PREDICATE THEME

3. **We ate lunch with them.**

Syntactic analysis:

[_S [_{NP} We] [_{VP} [_V ate] [_N lunch] [_{PP} [_P with] [_{NP} them]]]]

Semantic analysis:

We ate lunch with them
AGENT PREDICATE THEME AGENT

4. **He returned the book.**

Syntactic analysis:

[_S [_{NP} He] [_{VP} [_V returned] [_{NP} [_{Det} the] [_N book]]]]

Semantic analysis:

He returned the book
AGENT PREDICATE THEME

5. **John sang for everyone.**

Syntactic analysis:

[_S [_{NP} John] [_{VP} [_V sang] [_{PP} [_P for] [_{NP} everyone]]]]

Semantic analysis:

John sang for everyone
AGENT PREDICATE BENEFICIARY

The differences between syntactic and semantic parsing are:

- In syntactic parsing, every word gets assigned a part of speech, while in semantic parsing, some words may be grouped together.

- Syntactic parsing yields a tree structure of the sentence, while semantic role labelling is linear.
- No real-world empirical knowledge is needed to judge the syntactic parse of a sentence; only parts of speech are important. Semantic role labelling, on the other hand, depends more on the actual occurrence rather than the words used to describe it.

Question 3

Text:

നിശ്ചലചിത്രങ്ങളുടെ ഒരു ശ്രേണിയെ വളരെ പെട്ടെന്നു മാറ്റി മാറ്റി കാണിക്കുന്നതു വഴി ചലിക്കുന്ന ചിത്രങ്ങൾ സൃഷ്ടിക്കുന്ന ദൃശ്യ കലാരൂപമാണ് ചലച്ചിത്രം. ക്യാമറ ഉപയോഗിച്ച് ദൃശ്യങ്ങൾ പകർത്തിയൊ, ചിത്രങ്ങൾ അനിമേഷൻ ചെയ്തൊ മറ്റ് നൂതന സാങ്കേതിക വിദ്യകൾ ഉപയോഗിച്ചൊ ചലച്ചിത്രങ്ങൾ സൃഷ്ടിക്കാം.

ചലച്ചിത്രങ്ങൾ അവ നിർമ്മിക്കപ്പെടുന്ന സമൂഹങ്ങളുടെ സാംസ്കാരിക പ്രതിഫലനമാണ്. അതുപോലെ തന്നെ അവ തിരിച്ചും സമൂഹത്തിൽ സ്വാധീനം ചെലുത്തുന്നു. ചലച്ചിത്രങ്ങളെ പ്രധാനപ്പെട്ട ഒരു കലാരൂപമായും ജനപ്രിയ വിനോദോപാധിയായും കണക്കാക്കപ്പെടുന്നു. വിദ്യാഭ്യാസ ആവശ്യങ്ങൾക്കും ആശയങ്ങളുടെയും ആദർശങ്ങളുടെയും വ്യാപനത്തിനും ഇവ ഉപയോഗിക്കുന്നു. ചലച്ചിത്രങ്ങളുടെ ദൃശ്യഭാഷ അവയ്ക്ക് ഒരു സാർവ്വലോക വിനിമയശക്തി നൽകുന്നു. ചില ചലച്ചിത്രങ്ങൾ സംഭാഷണങ്ങൾ മറ്റ് ഭാഷകളിലേക്കു തർജ്ജമ ചെയ്തു അന്തർദ്ദേശീയ തലത്തിൽ പ്രശസ്തമായിട്ടുണ്ട്.

ചലച്ചിത്രങ്ങൾ നിശ്ചലചിത്രങ്ങളുടെ ഒരു ശ്രേണി ഉപയോഗിച്ചാണു ഉണ്ടാക്കുന്നത്. നിശ്ചലചിത്രങ്ങൾ അതിവേഗത്തിൽ തുടർച്ചയായി കാണിക്കുമ്പോൾ അവ ചലിക്കുന്നതായി തോന്നുന്നു. ഒരു ചിത്രം മാറ്റിയിട്ടും ഏതാനും നിമിഷാർദ്ധ നേരത്തേക്ക് അത് അവിടെ തന്നെ ഉള്ളതായി പ്രേക്ഷകനു തോന്നുകയും, അത് കാരണം ചിത്രങ്ങൾ തമ്മിലുള്ള ഇടവേള അറിയാതാവുകയും അങ്ങനെ ചിത്രങ്ങൾ ചലിക്കുന്നതായി തോന്നുകയും ചെയ്യുന്നു. കണ്ണിന്റെ സവിശേഷതയായ പെർസിസ്റ്റൻസ് ഓഫ് വിഷൻ ആണ് ഇതിനായി ഉപയോഗപ്പെടുത്തുന്നത്. ഒരു വസ്തുവിന്റെ പ്രതിബിംബം പതിനാറിൽ ഒരു സെക്കൻഡ് സമയത്തേക്ക് തങ്ങി നിൽക്കുന്നു. ആയതിനാൽ നിശ്ചല ദൃശ്യങ്ങൾ ഈ സമയത്തിൽ കൂടുതൽ വേഗത്തിൽ ചലിപ്പിച്ചാൽ ആ വസ്തു ചലിക്കുന്നതായി തോന്നുന്നു. ചലിക്കുന്ന ചിത്രത്തിൽ നിന്നാണു "ചലച്ചിത്രം" എന്ന പേരു രൂപപ്പെട്ടത്. സംസാര ഭാഷയിൽ ചിത്രം, പടം മുതലായ വാക്കുകളും ചലച്ചിത്രത്തെ സൂചിപ്പിക്കാനായി ഉപയോഗിക്കുന്നു. ഇംഗ്ലീഷ് വാക്കുകളായ ഫിലിം, മൂവി എന്നിവയും ഉപയോഗിക്കാറുണ്ട്. എന്നിരുന്നാലും "സിനിമ" എന്ന ഇംഗ്ലീഷ് വാക്കാണ് ഏറ്റവും അധികമായി ഉപയോഗിക്കപ്പെടുന്നത്.

Text (WX Transcription):

```
niScalaciwraffalYuteV oVru SreNiyeV valYareV peVtteVnnu mArYrYi mArYrYi
kANikkunnawu valYYi calikkunna ciwraffalY sqRtikkunna xqSyA kalArUpamAN
calacciwraM. kyAmarYa upayogicc xqSyaffalY pakarwwiyov, ciwraffalY animeRan
ceVywoV marYrY nUwana sAfkwika vixyakalY upayogiccoV calacciwraffalY
sqRtikkAM.
calacciwraffalY ava nirmmikkappeVtunna samUhaffalYuteV sAMskkArika
prawiPalanamAN. awupoleV wanneV ava wiriccuM samUhawwil svAXInaM ceVluwwunnu.
calacciwraffalYeV praXAnappeVtta oVru kalArUpamAyuM janapriya vinoxopAXiyAyuM
kaNakAkkippeVtunnu. vixyABYAsa AvaSyaffalYkkuM ASayaffalYuteVyuM
AxarSaffalYuteVyuM vyApanawwinuM iva upayogikkunnu. calacciwraffalYuteV
```

xqSyABARa avaykk oVru sArvvaloka vinimayaSakwi nalkunnu. cila calacciwraffalY saMBARaNaffalY marYrY BARakalYilekku warjjama ceVywu anwarxeSIya walawwil praSaswamAyittuNt. calacciwraffalY niScalaciwraffalYuteV oVru SreNi upayogiccANu uNtAkkunnaw. niScalaciwraffalY awivegawwil wutarccayAyi kANikkumpolY ava calikkunnawAyi wonnunu. oVru ciwraM mArYrYiyittuM ewAnuM nimirArxXa nerawwekk aw aviteV wanneV ulYlYawAyi prekRakanu wonnukayuM, aw kAraNaM ciwraffalY wammilulYlYa itavelYa arYiyAwAvukayuM affaneV ciwraffalY calikkunnawAyi wonnukayuM ceVyyunnu. kaNNinrYeV saviSeRawayAya peVrsisrYrYans oP viRan AN iwinAyi upayogappeVtuwwunnaw. oVru waswuvinrYeV prawibiMbAM pawinArYil oVru seVkkand samayawwekk waffi nilkunnu. AyawinAl niScala xqSyaffalY I samayawwil kUtuwal vegawwil calippicAl A wasw calikkunnawAyi wonnunu. calikkunna ciwrawwil ninnANu "calacciwraM" eVnna peru rUpappeVttaw. saMsAra BARayil ciwraM, pataM muwalAya vAkkukalYuM calacciwrawweV sUcippikkAnAyi upayogikkunnu. iMglIR vAkkukalYaya PiliM, mUvi eVnnivayuM upayogikkArYuNt. eVnnirunnAluM "sinima" eVnna iMglIR vAkkAN erYrYavuM aXikamAyi upayogikkuppeVtunnaw.

Annotation:

Word	Translation	Tag
niScalaciwraffalYuteV	of still images	A8 C1 M1-
oVru	a	Z99
SreNiyeV	sequence	N4
valYareV	very	A13.3
peVtteVnnu	quickly	N3.8
mArYrYi-mArYrYi	by changing	A2.1
kANikkunnawu	showing	X3.4 A10+
valYYi	by means of	X4.2
calikkunna	moving	M1+
ciwraffalY	images	A8 C1
sqRtikkunna	creating	A.1.1.1
xqSyA	visual	X3.4
kalArUpamAN	is an art form	C1 A4.1
calacciwraM	cinema	K4
kyAmarYa	camera	C1 X3.4
upayogicc	by using	A1.5.1+
xqSyaffalY	visuals	O2 X3.4
pakarwuiyoV	by copying, or	A1.1.1
ciwraffalY	images	A8 C1
animeRan	animation	A2.1 M1
ceVywoV	by doing, or	A1.1.1
marYrY	other	A1.8
nUwana	modern	T1.1.2 T3
sAfkewika	technological	Y1
vixyakalY	techniques	X4.2

Word	Translation	Tag
upayogiccoV	by using	A1.5.1+
calacciwraffalY	movies	X3.4 K4
sqRtikkAM	can create	A1.1.1 X9
calacciwraffalY	movies	X3.4 K4
ava	they	Z8
nirmmikkappeVtunna	being created	A1.1.1 A3
samUhaffalYuteV	groups	A4.1 S5
sAMskkArika	cultural	S1.1
prawipalanamAN	reflection	X3.4 W2 A10+
awupoleV	like that	A6.1+
wanneV	just	A14
ava	they	Z8
wiriccuM	back/again	M6
samUhawwil	society	S5
svAXInaM	influence	A2.1
ceVluwwunnu	exert	A1.1.1
calacciwraffalYeV	movies	X3.4 K4
praXAnappeVtta	main	A11.1+
oVru	a	Z99
kalArUpamAyuM	art-form and	C1 A4.1
janapriya	popular	E2 S2
vinoxopAXiyAyuM	means of entertainment	K1 X4.2
kaNakkAkkappeVtunnu	is considered	X2.1
vixyAByAsa	educational	P1
AvaSyaffalYkkuM	needs	X7
ASayaffalYuteVyuM	thoughts	X2.1
AxarSaffalYuteVyuM	ideas	X2.1
vyApanawwinuM	for the spreading	M1
iva	these	Z8
upayogikkunnu	are used	A1.5+
calacciwraffalYuteV	of movies	A9 X3.4 K4
xqSyaBARa	visual language	Q1 X3.4
avaykk	to them	Z8
oVru	a	Z99
sArvvaloka	universal	W1
vinimayaSakwi	currency	E2
nalkunnu	gives	A9+
cila	some	N5 A13.1
calacciwraffalY	movies	X3.4 K4
saMBARaNaffalY	dialogues	Q2.1
marYrY	other	A1.8
BARakalYilekku	into languages	Q3 M6
warjjama	translation	A2.1
ceVywu	having done	T1.1.1 A1.1.1

Word	Translation	Tag
anwarxeSIya	international	W1
walawwil	platform	M7
praSaswamAyittuNt	have become famous	S2
calacciwraffalY	movies	X3.4 K4
niScalaciwraffalYuteV	of still images	A8 C1 M1-
oVru	a	Z99
SreNi	sequence	N4
upayogiccANu	it is by using	A1.5.1
uNtAkkunnaw	that it is made	A1.1.1
niScalaciwraffalY	still images	A8 C1 M1-
awivegawwil	high speed	A13.3+ N8
wutarccayAyi	continuously	N6
kANikkumpoLY	when showing	X3.4
ava	they	Z8
calikkunnawAyi	to be moving	M1
wonnunnu	appear	A8
oVru	a	Z99
ciwraM	picture	C1
mArYrYiyittuM	having changed	A2.1
ewAnuM	approximately	A13.4
nimiRArxXa	half a second	T1.2 T1.3
nerawwekk	for a time	T1
aw	it	Z8
aviteV	there	M7
wanneV	just	A14
ulYlYawAyi	to be	A3
prekRakanu	to the viewer	S2 X3.4
wonnukayuM	appear, and	A8
aw	that	Z8
kAraNaM	because of	A2.2
ciwraffalY	images	A8 C1
wammilulYlYa	between each other	M7
itavelYa	interval	T1.3
arYiyAwAvukayuM	will not be noticed	X3.4 X2.5 Z6
affaneV	like that	A6.1+
ciwraffalY	images	C1
calikkunnawAyi	to be moving	M1
wonnukayuM	appear	A8
ceVyyunnu	do	A1.1.1
kaNNinrYeV	of the eye	B1 X3.4
saviSeRawayAya	peculiarity	A6.2- A4.2
peVrsisrYrYans oP viRan	persistence of vision	X3.4
AN	is	A3
iwinAyi	for this	A1.5.1

Word	Translation	Tag
upayogappeVtuwwunnaw	being used	A1.5.2
oVru	a	Z99
vaswuvinrYeV	of a thing	A9 02
prawibiMbaM	image	X3.4
pawinArYil oVru	one-sixteenth	N5
seVkkand	second	T1.3
samayawwekk	time	T1
waffi nilkkunnu	stays	M1-
AyawinAl	therefore	Z5
niScala	still	M1-
xqSyaffalY	images	X3.4
I	this	Z8
samayawwil	in time	T1.3
kUtuwal	more	N5+ A6
vegawwil	speed	N8
calippiccAl	if run	A2.2 M1 Z7
A	that	Z8
vaswu	thing	02
calikkunnawAyi	to be moving	M1
wonnunnu	appear	A8
calikkunna	moving	M1
ciwrawwil	in picture	X3.4 K4 C1
ninnANu	it is from	Z5
"calacciwraM"	"movie"	K4
eVnna	quot. particle	Z5
peru	name	Q3
rUpappeVttaw	that it was formed	A3
saMsAra	colloquial	Q3
BARayil	language	Q3
ciwraM	picture	X3.4 C1
pataM	picture	X3.4 C1
muwalAya	etc.	A4.1
vAkkukalYuM	words	Q3
calacciwrawweV	movie	X3.4 K4
sUcippikkAnAyi	to indicate	X2
upayogikkunnu	are used	A1.5
iMglIR	English	Q3
vAkkukalYAya	words	Q3
PiliM	film	X3.4 K4
mUvi	movie	X3.4 K4
eVnnivayuM	these also	Z8
upayogikkArYuNt	are used	A1.5
eVnnirunnAluM	even so	Z8
"sinima"	cinema	X3.4 K4

Word	Translation	Tag
eVnna	quot. particle	Z8
iMglIR	English	Q3
vAkkAN	it is the word	Q3
erYrYavuM	most	A13.2
aXikamAyi	most	A13.2
upayogikkuppeVtunnaw	being used	A1.5

Question 4

The challenges faced in the above tagging were:

- Malayalam words are frequently joined together, *e.g.*, **praSaswamAyittuNt** = “have/has become famous”, which is in fact **praSaswam** (“famous”) + **AyittuNt** (“have/has become”). This makes tagging tricky.
- Conjunctions (specifically “and” and “or”) are expressed as bound morphemes in Malayalam, and therefore do not get a separate tag.
- The use of **ceyyuka** (“to do”) as a kind of auxiliary verb makes it difficult to assign a semantic tag to.
- The quotative particle **eVnn(a)** has no tag.