

XV UNL School

Geneva, July 21-25, 2014

Program

- ~~July 21st~~
 - Introduction
 - Normalization
- ~~July 22nd~~
 - Tokenization
- ~~July 23rd~~
 - UNLization
 - corpus_org.txt
 - ngrammar.txt
 - dic.txt
 - dgrammar.txt
 - morphology.txt
 - syntax.txt
 - semantics.txt
- July 24th
 - NLization
- July 25th
 - Evaluation & Discussion



Day #4

Morning

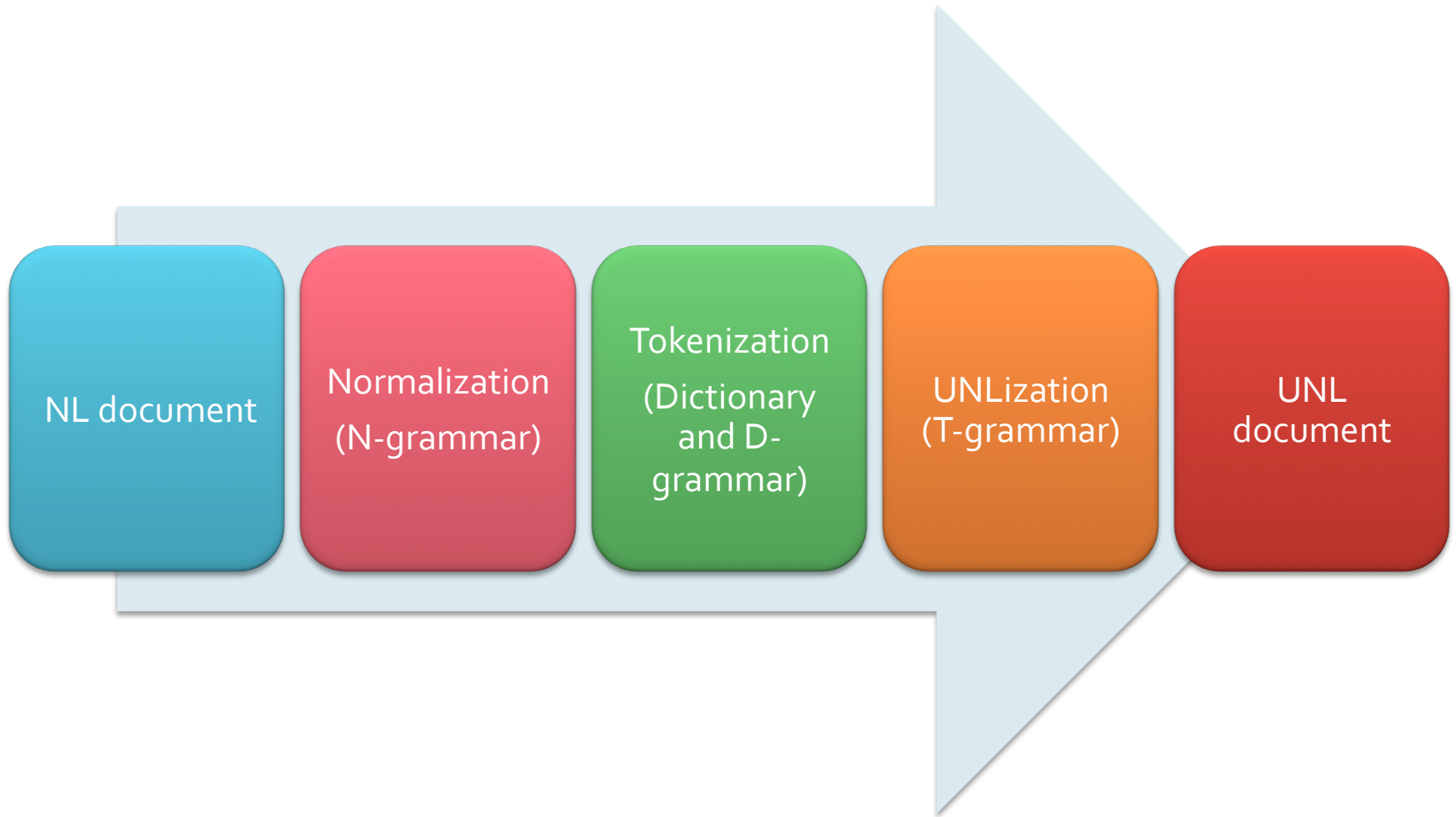
- UNLization

Lunch break

Afternoon

- NLization

UNLization

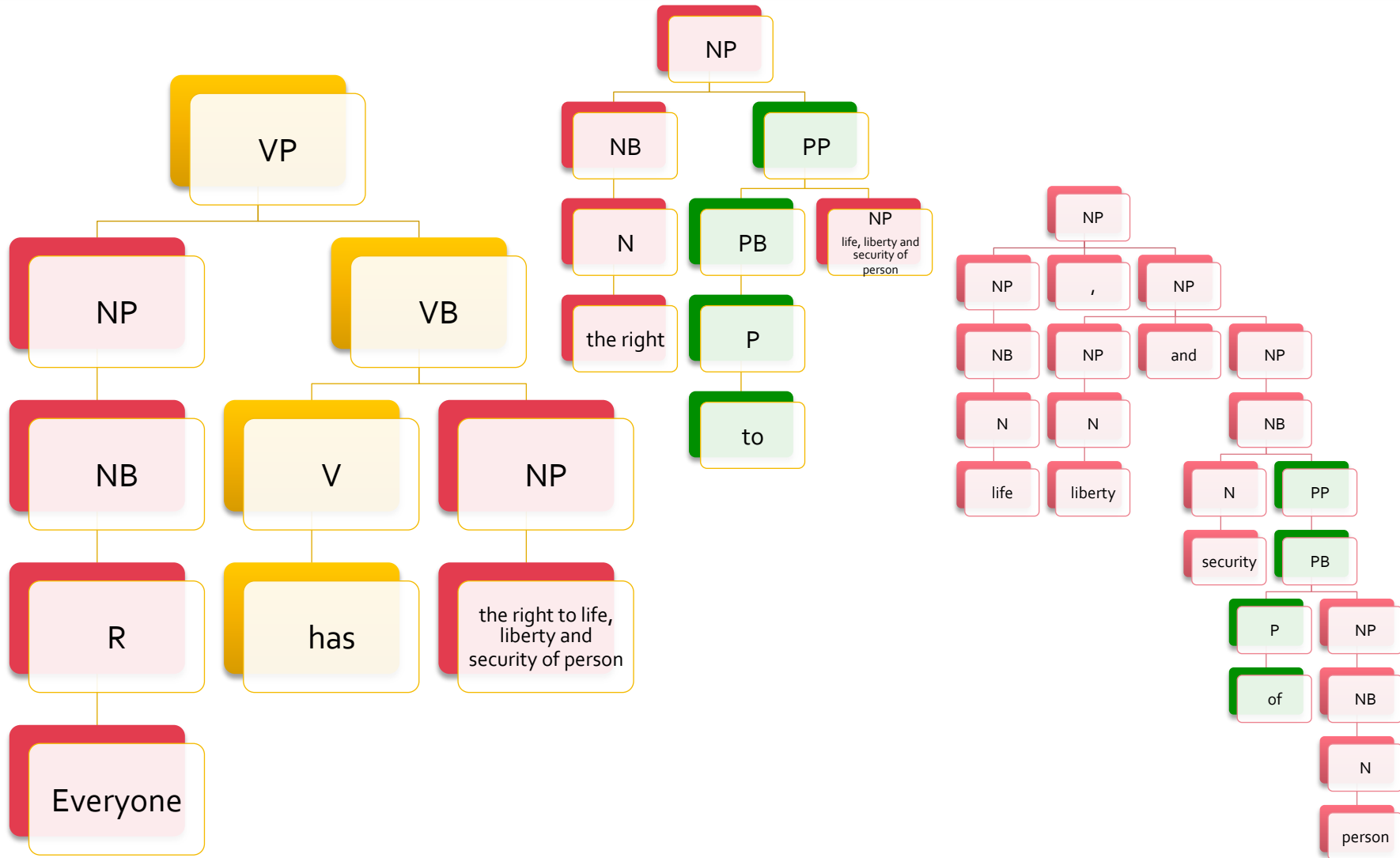


Exercise #11 (90 min)

- Goal: To UNLize, departing from your language, the following sentence:
 - Everyone has the right to life, liberty and security of person.
- Deliverable:
 - *Output*

Step #1

Draw the tree

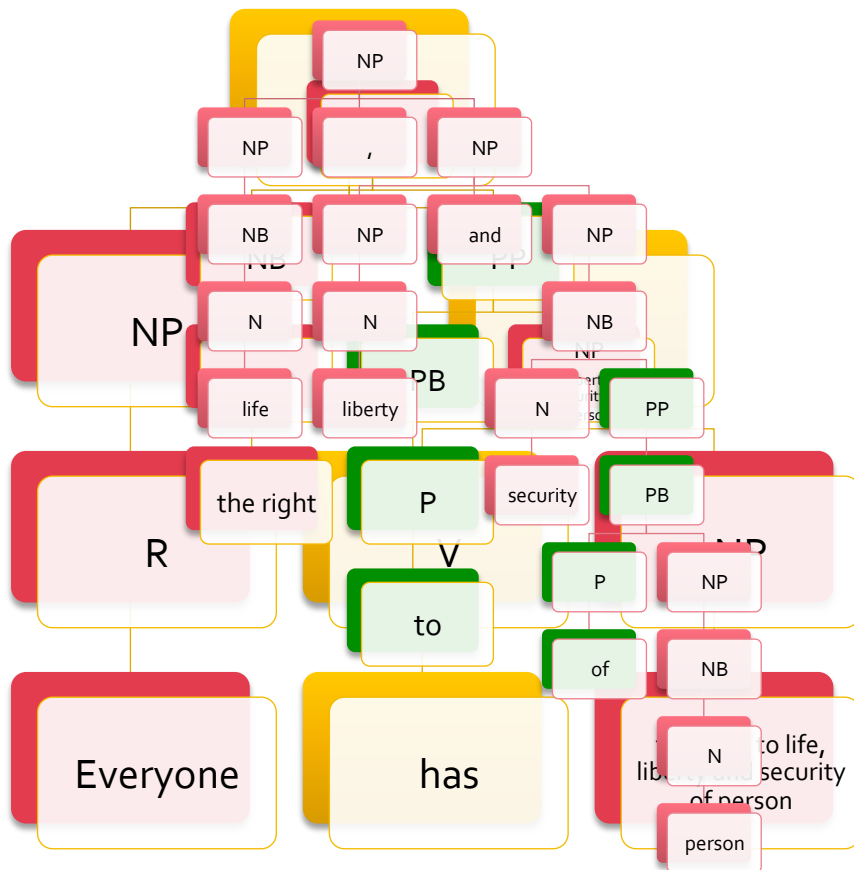


Steps

1. corpus_org.txt (>NL INPUT)
2. ngrammar.txt (>N-rules)
3. dic.txt (>Dictionary)
4. dgrammar.txt (>D-rules)
5. morphology.txt (>T-rules)
6. syntax.txt (>T-rules)
7. dsyntax.txt (>D-rules)
8. dearborisation.txt (>T-rules)
9. semantics.txt (>T-rules)

Step #5a

Extract the grammar



$(VP) := (NP)(VB);$
 $(NP) := (R);$
 $(VB) := (V)(NP);$
 $(NP) := (NB)(PP);$
 $(NB) := (N);$
 $(PP) := (PB)(NP);$
 $(PB) := (P);$
 $(NP) := (NP)(",")(NP);$
 $(NP) := (NP)("and")(NP);$
 $(NP) := (NB);$
 $(NB) := (N)(PP);$
 $(PP) := (PB);$
 $(PB) := (P)(NP);$

Step #5b

Reverse the order

(VP):=(NP)(VB);
(NP):=(R);
(VB):=(V)(NP);
(NP):=(NB)(PP);
(NB):=(N);
(PP):=(PB)(NP);
(PB):=(P);
(NP):=(NP)(",")(NP);
(NP):=(NP)("and")(NP);
(NP):=(NB);
(NB):=(N)(PP);
(PP):=(PB);
(PB):=(P)(NP);

(NP)(VB):=(VP);
(R):=(NP);
(V)(NP):=(VB);
(NB)(PP):=(NP);
(PB)(NP):=(PP);
(N):=(NB);
(P):=(PB);
(NP)(",")(NP):=(NP);
(NP)("and")(NP):=(NP);
(NB):=(NP);
(N)(PP):=(NB);
(PB):=(PP);
(P)(NP):=(PB);

Step #5c

Longest first

(NP)(VB):=(VP);
(R):=(NP);
(V)(NP):=(VB);
(NB)(PP):=(NP);
(PB)(NP):=(PP);
(N):=(NB);
(P):=(PB);
(NP)(",")(NP):=(NP);
(NP)("and")(NP):=(NP);
(NB):=(NP);
(N)(PP):=(NB);
(PB):=(PP);
(P)(NP):=(PB);

(NP)(",")(NP):=(NP);
(NP)("and")(NP):=(NP);

(NP)(VB):=(VP);
(NB)(PP):=(NP);
(PB)(NP):=(PP);
(V)(NP):=(VB);
(N)(PP):=(NB);
(P)(NP):=(PB);

(NB):=(NP);
(PB):=(PP);
(R):=(NP);

(N):=(NB);
(P):=(PB);

Step #5d

Indexation (syntax.txt)

BEFORE	AFTER
(NP)("")(NP):=(NP);	(NP,%x)("")(NP,%y):=(and(%x;%y),+LEX=N,+XP=NP);
(NP)("and")(NP):=(NP);	(NP,%x)("and")(NP,%y):=(and(%x;%y),+LEX=N,+XP=NP);
(NP)(VB):=(VP);	(NP,%x)(VB,%y):=(VP(%y;%x),+LEX=V,+XP=VP);
(NB)(PP):=(NP);	(NB,%x)(PP,%y):=(NP(%x;%y),+LEX=N,+XP=NP);
(PB)(NP):=(PP);	(PB,%x)(NP,%y):=(PP(%x;%y),+LEX=P,+XP=PP);
(V)(NP):=(VB);	(V,%x)(NP,%y):=(VB(%x;%y),+LEX=V,+XB=VB);
(N)(PP):=(NB);	(N,%x)(PP,%y):=(NB(%x;%y),+LEX=N,+XB=NB);
(P)(NP):=(PB);	(P,%x)(NP,%y):=(PB(%x;%y),+LEX=P,+XB=PB);
(NB):=(NP);	(NB,^XP):=(+XP=NP,+proj);
(PB):=(PP);	(PB,^XP):=(+XP=PP,+proj);
(R):=(NP);	(R,^XP,%x):=(+XP=NP);
(N):=(NB);	(N,^XB):=(+XB=NB,+proj);
(P):=(PB);	(P,^XB):=(+XB=PB,+proj);

Step #6

Disambiguation (dsyntax.txt)

- (proj)(^CCJ,^XP,^PUT,^STAIL)=0;
- (proj)({COMMA|CCJ})(^XP)=0;

Step #7

De-arborisation (dearborisation.txt)

```
[S:8]
{org}
Everyone has the right to life,
liberty and security of person
{/org}
{unl}
VP(:07, 00:01.@every.@person)
VB:07(have:03, :06)
NB:06(right:07.@def, :05)
PB:05("to":09, :04)
and:04(life:0B, :03)
and:03(liberty:0E, :02)
NB:02(security:0I, :01)
PB:01("of":0K, person:0M)
{/unl}
[/S]
```

```
[S:8]
{org}
Everyone has the right to life,
liberty and security of person
{/org}
{unl}
VS(:07, 00:01.@every.@person)
VC:07(have:03, :06)
NA:06(right:07.@def, :05)
PC:05("to":09, :04)
and:04(life:0B, :03)
and:03(liberty:0E, :02)
NA:02(security:0I, :01)
PC:01("of":0K, person:0M)
{/unl}
[/S]
```

Step #8

Interpretation (semantics.txt)

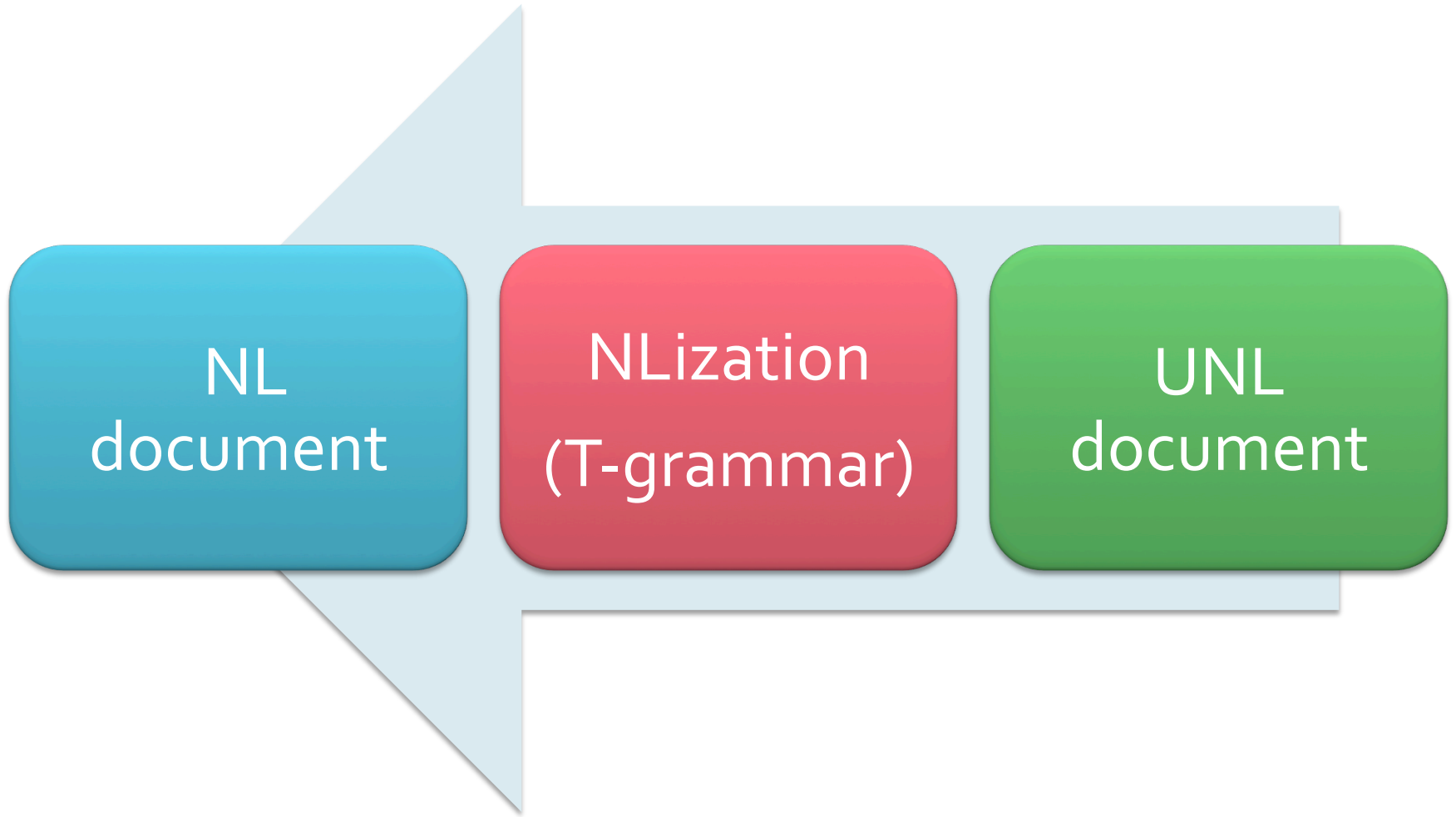
```
[S:8]
{org}
Everyone has the right to life,
liberty and security of person
{/org}
{unl}
VS(:07, 00:01.@every.@person)
VC:07(have:03, :06)
NA:06(right:07.@def, :05)
PC:05("to":09, :04)
and:04(life:0B, :03)
and:03(liberty:0E, :02)
NA:02(security:0I, :01)
PC:01("of":0K, person:0M)
{/unl}
[/S]
```

```
[S:8]
{org}
Everyone has the right to life,
liberty and security of person
{/org}
{unl}
exp(:07,
00:01.@every.@person)
cnt:07(have:03, :06)
cnt:06(right:07.@def, :04)
and:04(life:0B, :03)
and:03(liberty:0E, :02)
obj:02(security:0I, person:0M)
{/unl}
[/S]
```



Lunch Break

NLization



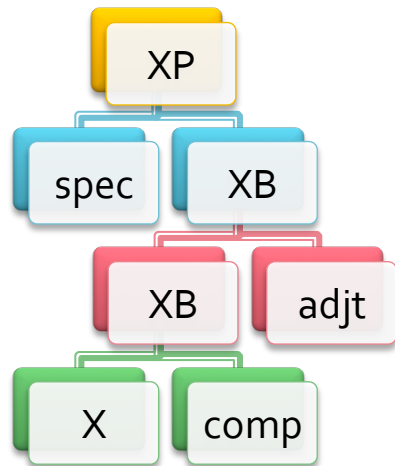
Steps

1. Syntax (syntax.txt)
2. Arborisation (arborisation.txt)
3. Linearization (linearisation.txt)
4. Morphological generation (morphology.txt)
5. Normalisation (normalisation.txt)

Syntactic Generation

Trees x Networks

TREE

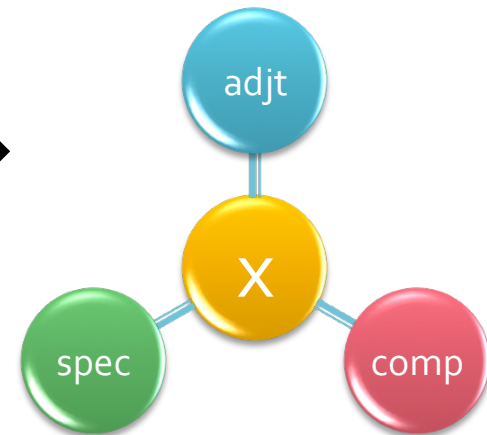


$XP(XB(XB(X;comp);adjt);spec)$

DEARBORISATION

ARBORISATION

NETWORK



$XS(X;spec)$
 $XA(X;adjt)$
 $XC(X;comp)$

Syntactic Generation (syntax.txt)

```
[S:8]
{org}
Everyone has the right to life,
liberty and security of person
{/org}
{unl}
exp(:07,
00:01.@every.@person)
cnt:07(have:03, :06)
cnt:06(right:07.@def, :04)
and:04(life:0B, :03)
and:03(liberty:0E, :02)
obj:02(security:0I, person:0M)
{/unl}
[/S]
```

```
[S:8]
{org}
Everyone has the right to life,
liberty and security of person
{/org}
{unl}
VS(:07, 00:01.@every.@person)
VC:07(have:03, :06)
NA:06(right:07.@def, :05)
PC:05("to":09, :04)
and:04(life:0B, :03)
and:03(liberty:0E, :02)
NA:02(security:0I, :01)
PC:01("of":0K, person:0M)
{/unl}
[/S]
```

Arborisation (arborisation.txt)

```
[S:8]
{org}
Everyone has the right to life,
liberty and security of person
{/org}
{unl}
VS(:07, 00:01.@every.@person)
VC:07(have:03, :06)
NA:06(right:07.@def, :05)
PC:05("to":09, :04)
and:04(life:0B, :03)
and:03(liberty:0E, :02)
NA:02(security:0I, :01)
PC:01("of":0K, person:0M)
{/unl}
[/S]
```

```
[S:8]
{org}
Everyone has the right to life,
liberty and security of person
{/org}
{unl}
VP(:07, 00:01.@every.@person)
VB:07(have:03, :06)
NB:06(right:07.@def, :05)
PB:05("to":09, :04)
and:04(life:0B, :03)
and:03(liberty:0E, :02)
NB:02(security:0I, :01)
PB:01("of":0K, person:0M)
{/unl}
[/S]
```

Linearization (linearisation.txt)

[S:8]
{org}
Everyone has the right to life,
liberty and security of person
{/org}
{unl}
VP(:07, 00:01.@every.@person)
VB:07(have:03, :06)
NB:06(right:07.@def, :05)
PB:05("to":09, :04)
and:04(life:0B, :03)
and:03(liberty:0E, :02)
NB:02(security:0I, :01)
PB:01("of":0K, person:0M)
{/unl}
[/S]

Everyone has right to life, liberty and
security of person

Morphological generation (morphology.txt)

everyone has right to life, liberty and
security of person

everyone has the right to life, liberty and
security of person

Normalisation (normalisation.txt)

Everyone has right to life, liberty and security of person

Everyone has the right to life, liberty and security of person.

Exercise #12 (90 min)

- Goal: To create the NLization grammar for sentences #1 and #8
- Deliverable:
 - Output



That's all, Folks!