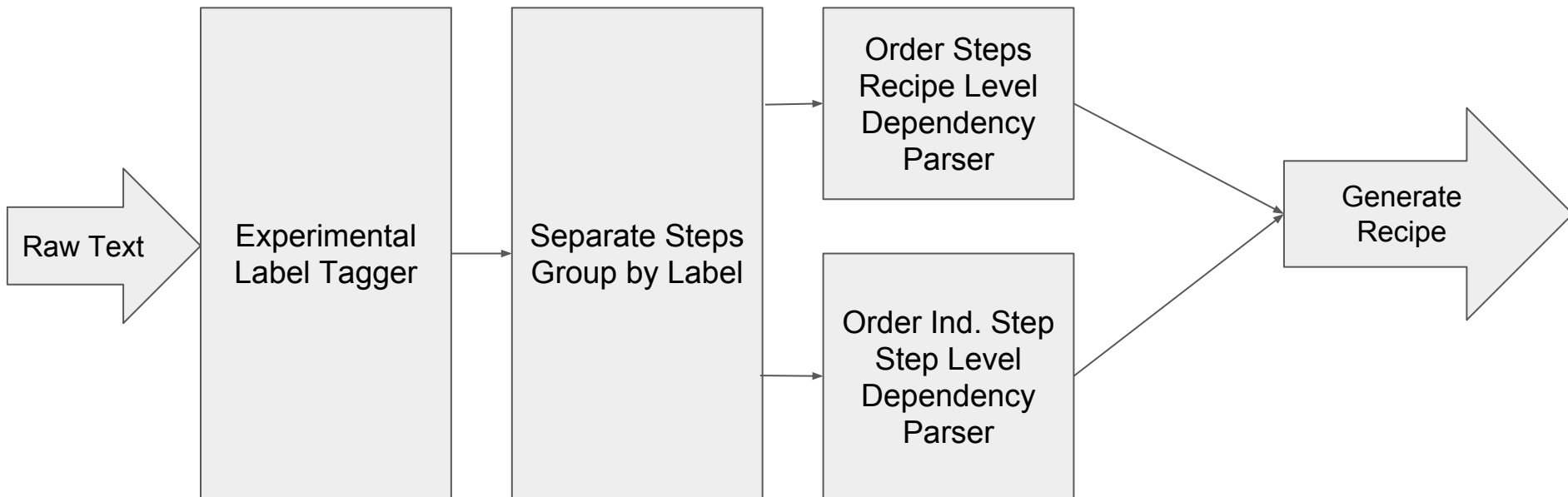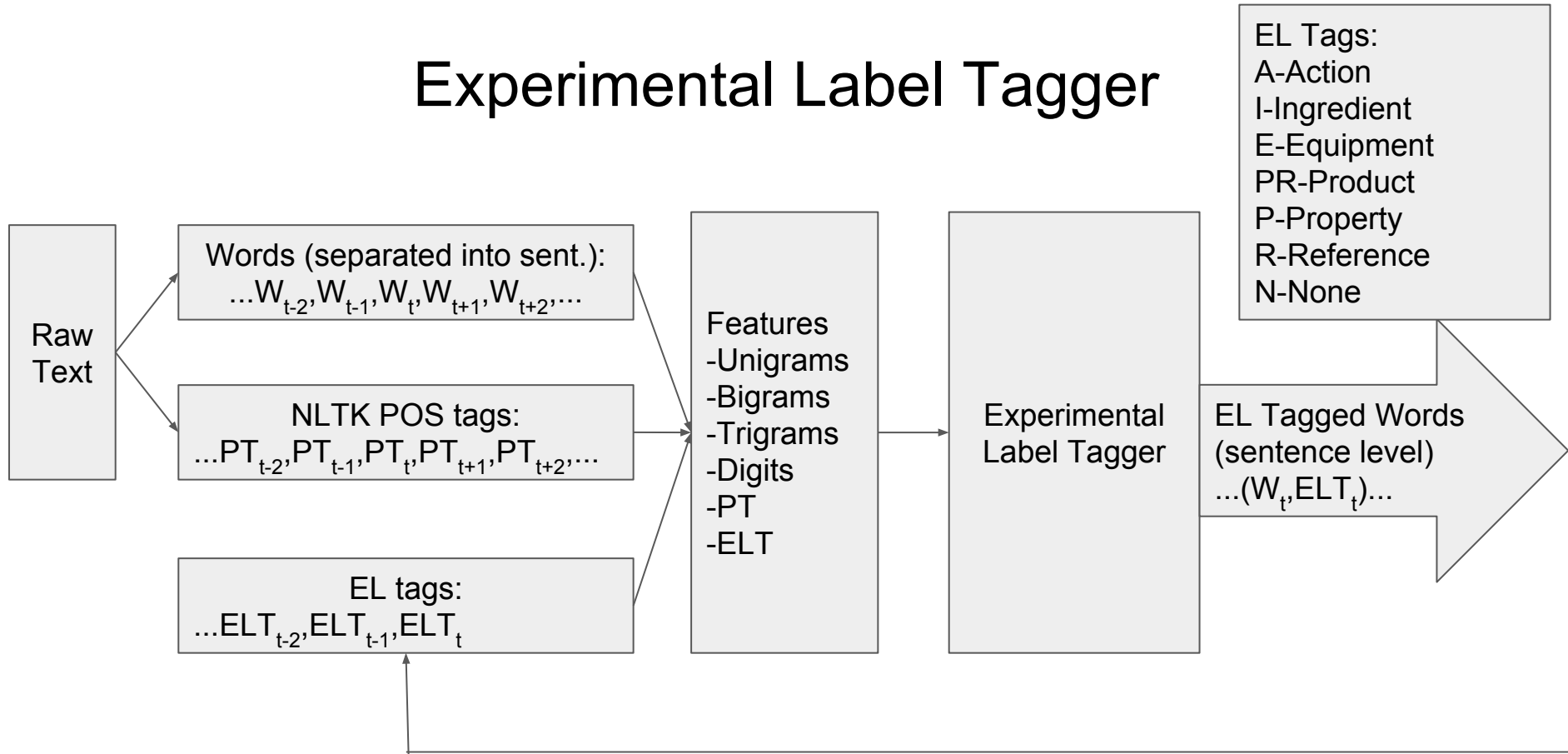# Turning Experimental Procedures into Machine-Readable Recipes

William Spitzer, Menghsuan Sam, Iveel Tsogsuren
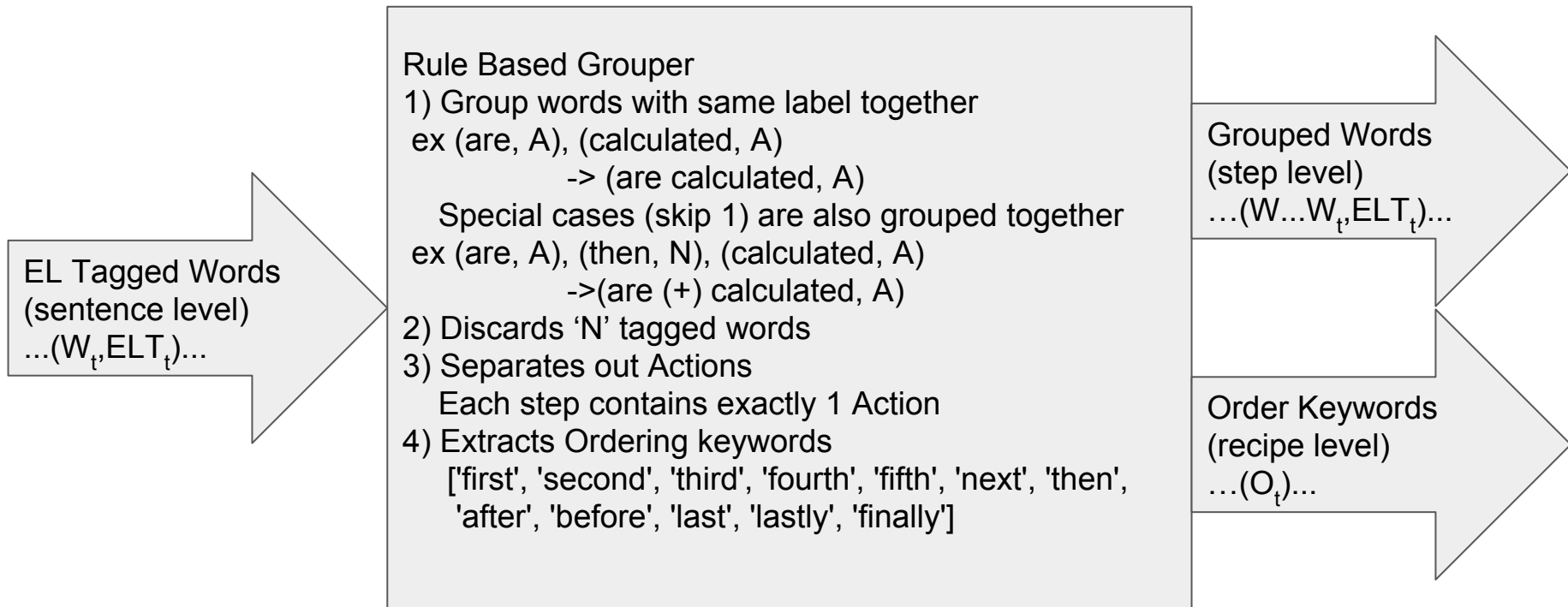
Process

Experimental Label Tagger

Raw Text

Words (separated into sent.):
...$W_{t-2}, W_{t-1}, W_t, W_{t+1}, W_{t+2}$,...

NLTK POS tags:
...$PT_{t-2}, PT_{t-1}, PT_t, PT_{t+1}, PT_{t+2}$,...

EL tags:
...$ELT_{t-2}, ELT_{t-1}, ELT_t$

Features
-Unigrams
-Bigrams
-Trigrams
-Digits
-PT
-ELT

Experimental Label Tagger

EL Tags:
A-Action
I-Ingredient
E-Equipment
PR-Product
P-Property
R-Reference
N-None

EL Tagged Words
(sentence level)
...(W_t, ELT_t)...

Experimental Label Grouper and Step Sorter

EL Tagged Words
(sentence level)
...(W_t, ELT_t)...

Rule Based Grouper
1) Group words with same label together
 ex (are, A), (calculated, A)
                -> (are calculated, A)
     Special cases (skip 1) are also grouped together
 ex (are, A), (then, N), (calculated, A)
                ->(are (+) calculated, A)
2) Discards 'N' tagged words
3) Separates out Actions
    Each step contains exactly 1 Action
4) Extracts Ordering keywords
    ['first', 'second', 'third', 'fourth', 'fifth', 'next', 'then',
     'after', 'before', 'last', 'lastly', 'finally']

Grouped Words
(step level)
...$(W...W_t, ELT_t)$...

Order Keywords
(recipe level)
...(O_t)...

Dependency Parser for Grouped Words

Grouped Words
(step level)
…(W...W_t,ELT_t)...

Features:
-Unigram
-Bigram
-Trigram
-EL tags

Transition Based
Dependency Parser
(uses SHIFT, LEFT,
RIGHT moves to
generate arcs)

List of Arcs (heads)
(step level)
…(H_t)...
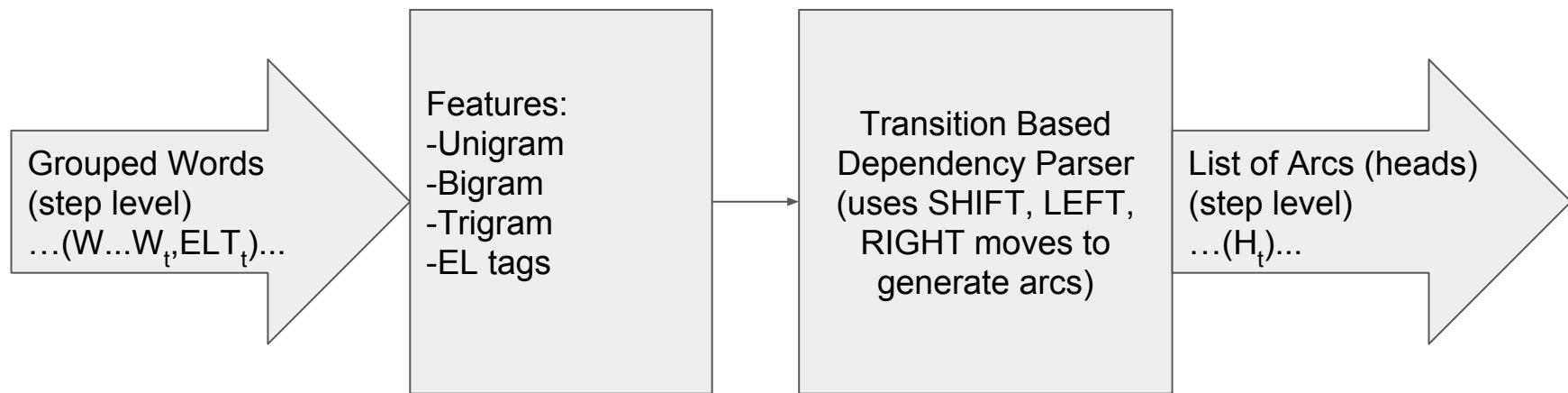
Dependency Parser for Recipe Steps

Order Keywords (recipe level) …(O_t)...

Features:
-Unigram
-Bigram
-Trigram

Transition Based Dependency Parser (uses SHIFT, LEFT, RIGHT moves to generate arcs)

List of Arcs (heads) (recipe level) …(H_t)...
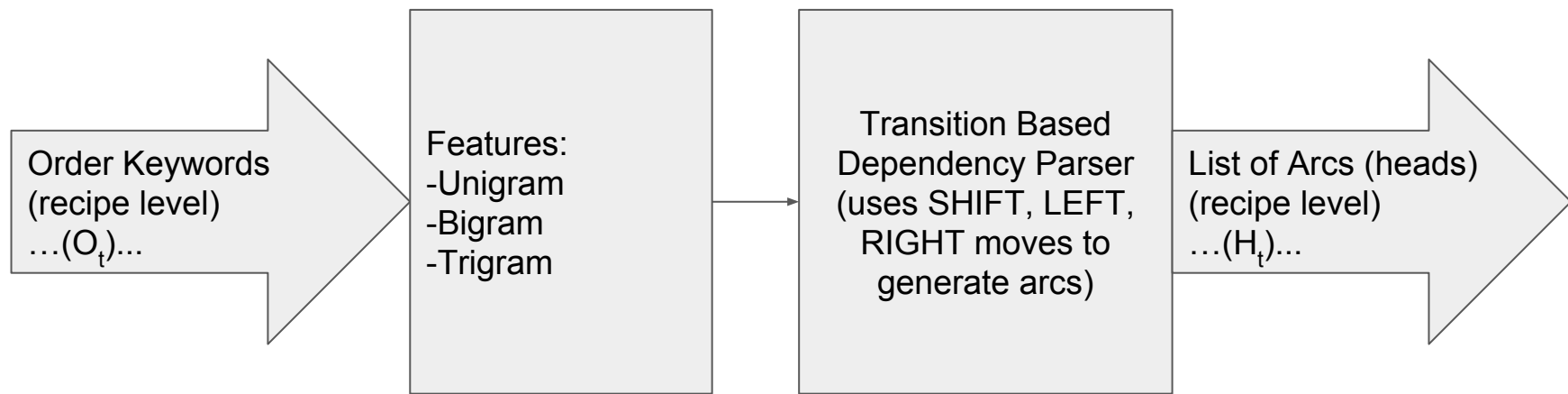
Generating Complete Recipe
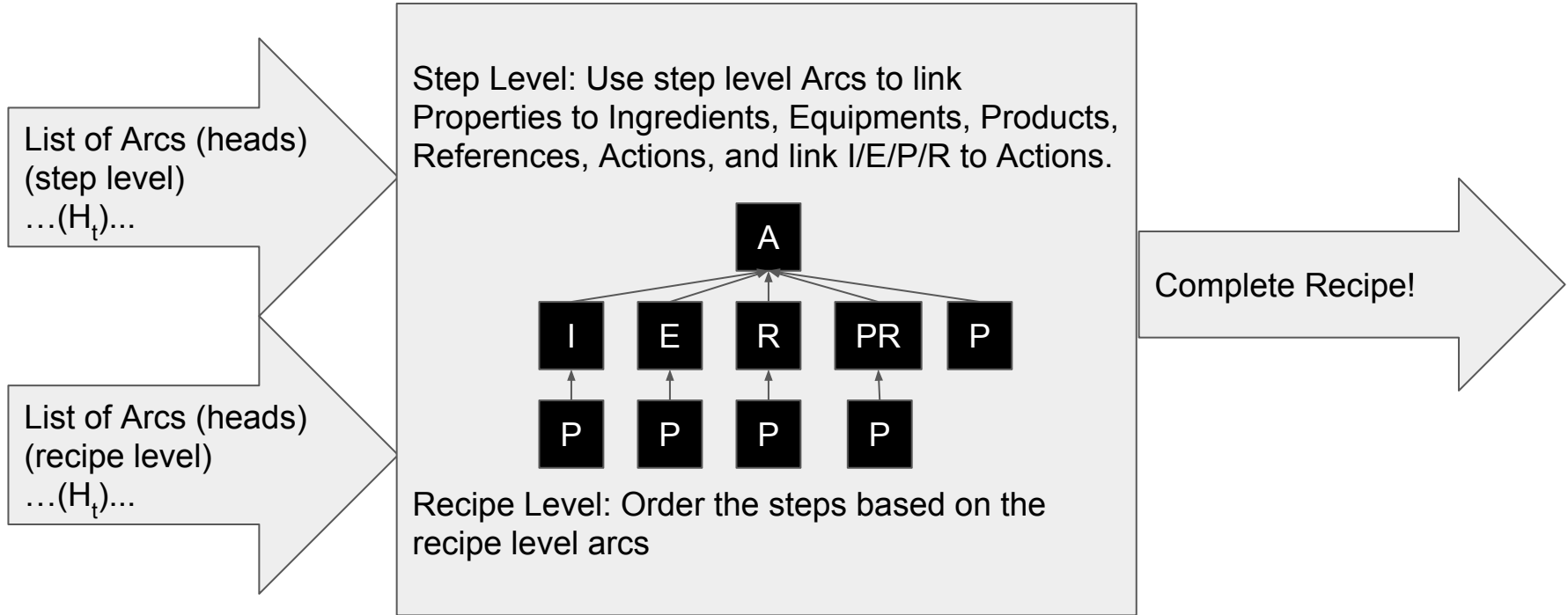
Sample Input/Output

The_N S/C_PR composites_PR were_A pulverized_A using_N a_N ball_E mill_E and_N sieved_A through_N a_N 25_P μm_P opening_P stainless_I steel_I sieve_I ._N
Slurries_PR were_A prepared_A using_N the_N S/C_R composites_R and_N a_N solution_N of_N 1_P wt_P %_P poly(vinylidene_N fluoride)_I (PVDF)_I in_N anhydrou
The_N slurries_R were_A applied_A to_N 10_P mm_P diameter_P aluminum_I current_N collectors_N and_N dried_A at_N 120_P C_P for_N 4_P h_P ._N
For_N the_N purpose_N of_N comparison_N ,_N the_N original_N mesoporous_I carbon_I (MPC)_I with_N 24.1_P wt_P %_P sulfur_I loading_N and_N WVA-1500_I (Mea
The_N batteries_PR were_A assembled_A as_N Swagelok_E cells_E by_N using_N the_N S/C_R composite_R coated_R aluminum_R foil_R (10_P mm_P diameter_P ,_N 7_
The_N organic_PR electrolytes_PR were_A solutions_N of_N Bis-_I (trifluoromethane)sulfonimide_I lithium_I (LiTFSI)_I (99.95%_P trace_P metals_P basis)_P i
The_N organic_R electrolyte_R filled_A the_N pores_P of_N the_N cathode_R and_N separator_R ._N
The_N cathode_R ,_N separator_R ,_N and_N anode_R were_A pressed_A by_N a_N spring_I to_N ensure_N tight_P contact_P ._N
A_N typical_N cell_N contained_N about_N 0.2_N mg_N of_N sulfur_N ._N
No_N excess_N of_N electrolyte_N was_N left_N in_N the_N assembled_N cell_N ._N
The_N batteries_R were_A tested_A in_N a_N Maccor_E 4000_E series_E battery_E tester_E ._N
The_N batteries_R were_A cycled_A between_P 1.0_P to_P 3.6_P V_P ._N
Each_N cycle_N was_N started_N with_N the_N discharge_N half-cycle_N ._N
Unless_N specified_N ,_N all_N batteries_R were_A tested_A at_N the_N same_N current_P of_P 0.5_P mA_P for_N both_N charging_P and_P discharging_P ._N
An_N average_N current_P density_P of_P 2.5_P A/g_P was_A used_A for_N all_N cells_R ._N
The_N cutoff_PR current_PR for_N the_N charge_E cycle_E was_A set_A to_N 0.05_P mA_P ._N
The_N calculation_N of_N specific_PR discharge_PR capacities_PR is_A based_A on_N the_N mass_P of_P elemental_P sulfur_P ._N

Evaluation		Averaged over 10 shuffled runs	
	EL Tagger	Step Level Dependency Parser	Recipe Level Dependency Parser
Independently	0.4744	0.8441	0.8807
Pipeline	0.474	0.4287	0.2322

```
Python 2.7.10 Shell
File  Edit  Shell  Debug  Options  Window  Help

Step 1:
{'A': [('were pulverized', {'E': [('S/C composites', {}), ('ball mill', {})]})]
}
Step 2:
{'A': [('sieved', {'P': [('25 m opening stainless steel sieve .', {})]})]}
Step 3:
{'A': [('were prepared', {'I': [('anhydrous N-methyl-2-pyrolidinone (NMP)', {'P
': ['1 wt %']}), 'P': [('ratio of 1:5', {}), 'E': [('S/C composites', {}), ('
solution', {})]})]}
Step 4:
{'A': [('were applied', {'PR': [('slurries', {})], 'P': [('10', {}), ('120 C (+
) 4 h', {})]})]}
Step 5:
{'A': [('prepared', {'P': [('24.1 wt % sulfur loading', {'P': ['25.2 wt % were'
]})], 'E': [('identical procedure', {})]})]}
Step 6:
{'A': [('was used', {'E': [('S/C', {})]})]}
Step 7:
{'A': [('were assembled', {'PR': [('batteries', {})], 'P': [('Swagelok', {}), (
'7', {}), ('10 mm diameter)', {})], 'E': [('S/C composite coated aluminum foil
(10 mm diameter', {}), ('lithium foil (7 mm thickness', {}), ('organic', {}),
'I': [('Celgard 3225 separator (10.3 mm diameter)', {})]})]}
Step 8:
{'A': [('were', {'PR': [('organic electrolytes', {})], 'I': [('(trifluoromethan
e)sulfonimide lithium (LiTFSI) (99.95% trace', {}, ('(DOL)', {})], 'P': [('55:
40', {})]})]}
Step 9:
{'PR': [('organic electrolyte', {'A': [('filled', {})]})]}
Step 10:
{'A': [('were pressed', {'E': [('spring', {})]})]}
Step 11:
{'A': [('were tested', {'PR': [('batteries', {})], 'E': [('Maccor 4000 series b
attery tester', {})]})]}
Step 12:
{'A': [('were cycled', {'PR': [('batteries', {})], 'P': [('between 1.0 to 3.6 V
', {})]})]}
Step 13:
{'A': [('was started', {})]}
Step 14:
{'A': [('were tested', {'P': [('0.5 mA', {})], 'R': [('batteries', {})]})]}
Step 15:
{'A': [('was set', {'P': [('0.05', {})]})]}
Step 16:
{'A': [('is', {'I': [('discharge', {}), ('sulfur', {})]})]}
>>>
```