

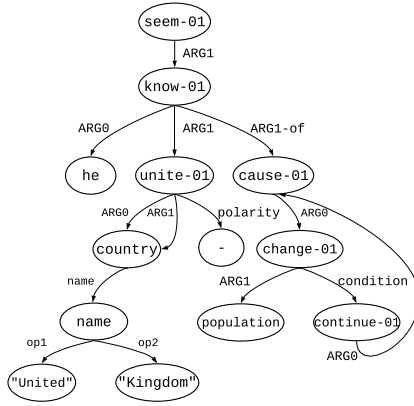
Supplementary Material For: Better Transition-Based AMR Parsing with a Refined Search Space

Zhijiang Guo and Wei Lu

Singapore University of Technology and Design

8 Somapah Road, Singapore, 487372

zhijiang-guo@mymail.sutd.edu.sg, luwei@sutd.edu.sg



It seems that he knows under the population changes
(if it continues) that United Kingdom will not unite.

Figure 1: An example AMR graph for a sentence.

1 AMR Concepts Re-categorization Details

In Section 3.1, we categorize AMR concepts into lexical and non-lexical. Here, we further categorize non-lexical concepts into these classes:

1. **Named Entity:** In AMR, any concept that have the concept name as child and name connects to their child concepts by relations starting with *op* can be named entity. An example is the concept *country* in Figure 1. Moreover, subgraph rooted by a named entity can be compressed into a new concept in compact AMR graph as concept *ENcountry* shown in Figure 2.
2. **Date Entity:** Similar to Named Entity, concept starts with *date*— such as *date-entity* and have connects their child concepts by relations *year*, *month*, *day*, etc. is classified into this category.
3. **Other Entity:** Applies to the concept ends with *-entity*, such as *url-entity*, which is invoked by a website in the input sentence.

4. **Quantity:** Applies to concept ends with *-quantity*, such as *distance-quantity*, which is invoked by the tokens indicate distance like *8 mile*.

5. **Interrogative:** Applies to the concept interrogative with an incoming mode relation arc, which is invoked by yes-no questions in the input sentence. An example is the token *whether* in the sentence.

6. **Imperative:** Applies to the concept imperative with an incoming mode relation arc, which is invoked by verbs with imperative in the input sentence. An example is token *go* in the sentence *Let's go*.

7. **Expressive:** Applies to concept expressive with an incoming mode relation arc, which is invoked by verbs with exclamation in the input sentence. An example is the token *went* in the sentence *We went!*

8. **Degree:** Applies to concept more or most with an incoming relation arc *degree*, which is invoked by comparative or superlative in the input sentence. An example is the token *better*.

Also, we provide the detailed list of non-lexical concepts used by the Hybrid Aligner.

Non-Lexical Concepts: person, man, aircraft, government-organization, political-party, product-of, multiple, criminal-organization, ethnic-group, religious-group, relative-position, railway-line, airport, club, ocean, sports-facility, show, aircraft-type, war, worm, channel, university, book, military, magazine, animal, peninsula, department, bank, car-make, hospital, city-district, natural-disaster, medicine, network, newspaper, valley, band, city, ship, palace, person, school, spaceship, treaty, operation, planet, movie, street,

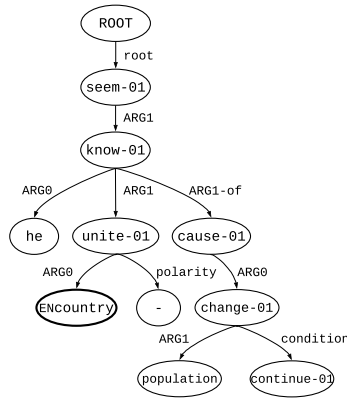


Figure 2: The compact AMR graph for the example in Figure 1.

district, company, bay, earthquake, journal, family, event, publication, language, republic, hotel, mountain, agency, moon, road, initiative, system, name, island, game, disease, station, lawyer, park, local-region, research-institute, country, doctor, worship-place, building, organization, seminar, music, river, rocket, dynasty, vehicle, county, incident, state, world-region, range, province, thing, country-region, law, broadcast-program, capital, continent, facility, revolution, conference, location, page, territory, man, sea, product, award, pass, destroyer, ethnic-group, son, plane, television, mission, drug, meet-03, government-organization, strait, political-party, missile, league, program, criminal-organization, group, project, religious-group, team, radio, museum, husband, festival, have-rel-role-91, have-org-role-91, be-located-at-91, have-subevent-91, have-purpose-91, have-polarity-91, be-temporally-at-91, be-destined-for-91, street-address-91, have-frequency-91, have-quant-91, byline-91, date-entity, molecular-physical-entity, ordinal-entity, percentage-entity, phone-number-entity, email-address-entity, rate-entity-91, score-entity, url-entity, monetary-quantity, distance-quantity, area-quantity, volume-quantity, temporal-quantity, frequency-quantity, speed-quantity, acceleration-quantity, mass-quantity, force-quantity, pressure-quantity, energy-quantity, power-quantity, charge-quantity, potential-quantity, resistance-quantity, inductance-quantity, magnetic-field-quantity, magnetic-flux-quantity, radiation-quantity, fuel-consumption-quantity, numerical-quantity, information-quantity, concentration-quantity, catalytic-activity-quantity, acidity-quantity, seismic-quantity, temperature-quantity, interrogative, imperative, expressive, more, most.

In order to use the Hybrid Aligner, we first linearize the AMR graph and remove all relations and non-lexical concepts as shown above. Then we run the IBM model. For non-lexical concepts, we can align them to the same span that its child concepts aligned to.

1. **Named Entity**: Matches a span that exactly matches its children with outgoing opn relation arcs in numerical order.
2. **Date Entity**: Matches any permutation of its children with outgoing day, month, year, season, time, timezone, century relation arcs.
3. **Other Entity**: Matches a span that exactly matches its child with an outgoing value relation arc.
4. **Quantity**: Matches a span that exactly matches its child with an outgoing unit relation arc.
5. **Interrogative**: Matches a span that exactly matches its parent with an incoming mode relation arc.
6. **Imperative**: Matches a span that exactly matches its parent with an incoming mode relation arc.
7. **Expressive**: Matches a span that exactly matches its parent with an incoming mode relation arc.
8. **Degree**: Matches a span that exactly matches its parent with an incoming degree relation arc.

For some specific lexical concepts, the IBM model may not yield optimal alignments. For these concepts, we also use a set of rules to greedily align them to spans similar to JAMR.

1. **Negative Polarity**: Applies to the concept $-$ with an incoming polarity relation arc, which indicates negation in the input sentence. An example is token *illegal* in the input sentence. The concept is aligned to a word beginning with “un”, “in”, “il”, etc.
2. **Positive Polarity**: Applies to the concept $+$ which is invoked by certain tokens including *please*, *could*.

Model Components	Hyper-Parameters
Pretrained Embeddings	100
POS Embeddings	12
Relation Embeddings	20
Action Embeddings	100
Predicate Embeddings	100
Entity Embeddings	100
Generated Concept Embeddings	100
LSTM Hidden Size	200
LSTM Layers	2
Dropout Rate	0.2
UNK Probability	0.2

Table 1: Model hyper-parameters.

3. **Wh-Question:** Applies to the concept `amr-unknown`, which is invoked by `wh`-questions in the input sentence, such as *what*, *why*, etc.

2 Model Parameters and Optimization Details

We selected hyper-parameters based on the best performance on the development set. Hyper parameters for the model are summarized in Table 1.

3 Constraints of the parser

We define **Acyclicity**, **Simple**, **Non-terminal restricted** and **Reentrancy restricted** for the compact AMR graph. Each property corresponds to certain constraints on the valid action set \mathcal{A} for each parsing state.

- **Acyclicity:** In order to maintain this property, we check whether a dependent being attached is the ancestor of the head that is attaching. Take concept `cause-01` in Figure 1 as an example. If `cause-01` is at the front of the buffer and concept `continue-01` is on the top of the stack, action `RIGHTLABEL` will be applied. However, `cause-01` is actually the ancestor of `continue-01`. Thus, `RIGHTLABEL` action is forbidden in the current state in order to maintain Acyclicity.
- **Simple:** In order to maintain this property, we check whether a dependent being attached has already been attached by the same head that is attaching. Take concept `country` in Figure 1 as an example. If concept `country` had been attached by concept unit with relation `ARG0`, and unit is on the top of the stack, action `RIGHTLABEL` would be forbidden under this circumstance.
- **Non-terminal restricted:** In order to maintain this property, certain concepts are for-

bidden to have child concepts, including: the child concepts of Name Entity concepts, such as `name` and its child concepts connected by relations start with `op`; the child concepts of Date Entity; the child concepts of Other Entities; child concepts of Quantity; Negative Polarity; Positive Polarity; Interrogative; Imperative; Expressive and Degree. An example is that when concept `—` is at the front of buffer, action `LEFTLABEL` is forbidden.

- **Reentrancy restricted:** In order to maintain this property, certain concepts are forbidden to be a reentrancy node, including: Name Entity and its child concepts; Date entity and its child concepts; Other Entity and its child concepts; Negative Polarity; Positive Polarity; Interrogative; Imperative; Expressive; Wh-Questions and Degree. An example is that when concept `—` which has already been attached as a dependent is at the front of buffer, action `RIGHTLABEL` is forbidden. Apart from this, we also set a maximum number (seven) of reentrancy nodes for all instances. If there exist seven reentrancy nodes in the growing compact AMR graph, no more reentrancy will be allowed.

Apart from these constraints, we also apply some basic constraints on different actions.

- **SWAP:** Repetitive `SWAP` actions are disallowed to avoid infinite swapping.
- **GEN:** The maximum number of consecutive `GEN` actions is set to three. Also, the maximum number of `GEN` actions for one instance is set to ten. We apply these constraints in order to prevent infinite `GEN` actions and oversize AMR graph.
- **MERGE:** The maximum number of consecutive `GEN` actions is set to eleven.

In order to impose these constraints during training phase, we remove certain relations for some AMR graphs in the training set. In the testing stage, all generated compact AMR graphs respect such properties since we have already imposed constraints on the valid action set for each state.