ABSTRACT MEANING REPRESENTATION LINTAS BAHASA BERDASARKAN MODEL GRAF PRALATIH

PROPOSAL TESIS

Karya tulis sebagai salah satu syarat
kelulusan MK IF5099 Metodologi Penelitian/Tesis 1

Oleh
MOCH. NAFKHAN ALZAMZAMI
NIM: 23522007
(Program Studi Magister Informatika)

INSTITUT TEKNOLOGI BANDUNG
Desember 2022

ABSTRACT MEANING REPRESENTATION LINTAS BAHASA BERDASARKAN MODEL GRAF PRALATIH

Oleh
Moch. Nafkhan Alzamzami
NIM: 23522007
(Program Studi Magister Informatika)

Institut Teknologi Bandung

Menyetujui
Calon Tim Pembimbing

Tanggal 15 Desember 2022

Calon Pembimbing,

(Dr. Masayu Leylia Khodra, S.T, M.T.)

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

DAFTAR SINGKATAN DAN LAMBANG

SINGKATAN	**Nama**	**Pemakaian pertama kali pada halaman**
1-NN	*1-Nearest Neighbour*	28
AMR	*abstract meaning representation*	ii
AMRBART	*Graph Pre-training for AMR Parsing and Generation*	ii, 4
BART	*Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension*	3
BERT	*Bidirectional Encoder Representations from Transformers*	14
DFS	*depth-first search*	21
GPT2	*Generative Pre-trained Transformer 2*	14
GRU	*gated recurrent neural network*	14
LABSE	*Language-agnostic BERT Sentence Embedding*	15
LSTM	*long short-term memory*	14
MLM	*Masked Language Modeling*	15
NLP	*natural language processing*	1
PLM	*pretrained language model*	20
RNN	*recurrent neural network*	14
seq2seq	*sequence-to-sequence*	1
SMATCH	*Evaluation Metric for Semantic Feature Structures*	ii, 2
SRL	*semantic role labeling*	1
stog	*Sequence-to-Graph Transduction*	3
TLM	*Translation Language Modeling*	15
XL-AMR	*Cross-lingual AMR*	2
XLM	*Cross-lingual Language Model*	iv, 15

Bab I Pendahuluan

I.1 Latar Belakang

Teks yang dapat dipahami oleh manusia tidak dapat dipahami secara langsung oleh model *natural language processing* (NLP) sehingga diperlukan sebuah bentuk representasi lain. Ada beberapa teknik dalam membuat representasi dari sebuah teks, seperti *bag of words*, *n-gram*, *word embedding*, *semantic role labeling* (SRL), dan AMR. Dari representasi-representasi tersebut, dapat dilakukan berbagai macam *task* seperti peringkasan teks, klasifikasi sentimen, mesin translasi, *question-answering*, deteksi parafrasa, dan lain-lain.

AMR merupakan salah satu representasi yang memperhatikan semantik pada tingkat kalimat dalam bentuk struktur data graf berarah yang mempunyai akar (Banarescu dkk., 2013). AMR awalnya didesain untuk merepresentasikan kalimat berbahasa Inggris sehingga bentuk graf AMR juga dituliskan dengan Bahasa Inggris. Sebuah teks yang mengandung banyak kalimat dapat direpresentasikan menjadi beberapa graf AMR, dengan setiap graf merepresentasikan setiap kalimat. Dalam membuat sebuah AMR dari sebuah kalimat, perlu dilakukan sebuah proses yang dinamakan AMR *parsing*.

AMR *parsing* dapat dikategorikan menjadi dua pendekatan, yaitu dua tahap *parsing* dan satu tahap *parsing* (Cai dan Lam, 2020). Pada pendekatan dua tahap *parsing* digunakan desain *pipeline* untuk identifikasi konsep dan prediksi relasi. Pada pendekatan satu tahap *parsing* dikategorikan menjadi tiga jenis, yaitu *parsing* berbasis transisi, *parsing* berbasis *sequence-to-sequence* (seq2seq), dan *parsing* berbasis graf. Pendekatan satu tahap jenis *parsing* berbasis transisi dilakukan dengan memproses kalimat dari kiri ke kanan dan membangun grafik secara bertahap dengan secara bergantian memasukkan simpul atau sisi baru. Pendekatan satu tahap jenis *parsing* berbasis seq2seq dengan melihat *parsing* sebagai transduksi urutan linear ke urutan linear juga dengan memanfaatkan linearisasi grafik AMR. Pendekatan satu tahap jenis *parsing* berbasis graf di mana setiap langkah waktu, simpul baru beserta koneksinya ke simpul yang ada diputuskan bersama secara

berurutan maupun paralel.

Dalam mengukur kelayakan hasil graf AMR yang dihasilkan dari suatu teknik, digunakan metrik *Evaluation Metric for Semantic Feature Structures* (SMATCH) (Cai dan Knight, 2013). SMATCH mengukur derajat *overlap* antara dua struktur fitur semantik graf AMR. Teknik-teknik AMR *parsing* tersebut diukur kelayakannya menggunakan dataset pasangan teks berbahasa Inggris dengan graf AMR-nya.

Pada AMR berbahasa Indonesia, telah dikembangkan untuk *parsing* berbasis aturan (Severina dan Khodra, 2019), berbasis *dependency parser* (Ilmy dan Khodra, 2020), dan berbasis *cross-lingual* (Putra, 2022). Teknik AMR *parsing* berbasis aturan dan *dependency parser* masih terbatas karena kurangnya panduan anotasi dan PropBank untuk adaptasi AMR dalam Bahasa Indonesia. Dataset yang digunakan sebagai data pelatihan teknik tersebut terdiri dari 1,130 pasang kalimat dan AMR berbahasa Indonesia. Anotasi pada dataset tersebut hanya terbatas pada 6 relasi saja, yaitu `:ARG0`, `:ARG1`, `:name`, `:time`, `:location`, `:mod` (Ilmy dan Khodra, 2020). Konsep yang membutuhkan argumen relasi lebih banyak yang tidak tersedia akan diabaikan sehingga AMR Bahasa Indonesia tidak dapat merepresentasikan kalimat yang lebih panjang dan kompleks.

AMR *parsing* awalnya digunakan untuk mengubah dari kalimat berbahasa Inggris menjadi sebuah graf AMR berbahasa Inggris. AMR dalam bahasa selain Bahasa Inggris memiliki banyak limitasi karena AMR berbahasa Inggris sudah dikembangkan lebih dahulu dan memiliki anotasi konsep dan relasi lengkap yang tidak dimiliki AMR bahasa lain. Ada beberapa teknik yang dapat membaca kalimat bahasa lain menjadi graf AMR berbahasa Inggris untuk merepresentasikan kalimat dari bahasa selain Bahasa Inggris. Damonte dan Cohen (2018) mengusulkan teknik AMR *parsing* lintas bahasa, yaitu teknik mengubah kalimat dari bahasa selain Bahasa Inggris menjadi AMR berbahasa Inggris. Dalam teknik tersebut, dibutuhkan dataset pasangan teks berbahasa selain Bahasa Inggris yang dituju dengan graf AMR-nya. Model *Cross-lingual AMR* (XL-AMR) memiliki kinerja SMATCH 53.0 untuk Bahasa Cina, 53.0 untuk Bahasa Jernam, 58.1 untuk Bahasa Italia, dan 58.0 untuk Bahasa Spanyol.

Pada Bahasa Indonesia, beberapa teknik AMR *parsing* untuk teks berbahasa Indonesia juga dikembangkan. Beberapa AMR *parsing* Bahasa Indonesia yang telah dikembangkan adalah AMR *parsing* berbasis aturan (Severina dan Khodra, 2019), berbasis *dependency parser* (Ilmy dan Khodra, 2020), dan berbasis lintas bahasa (Putra, 2022). Teknik AMR *parsing* oleh Putra (2022) menggunakan teknik *training* XL-AMR (Blloshmi dkk., 2020) dengan model *Sequence-to-Graph Transduction* (stog) (Zhang dkk., 2019) untuk melakukan AMR *parsing* dari kalimat berbahasa Indonesia menjadi graf AMR berbahasa Inggris. Model stog menggunakan pendekatan dua tahap *parsing*, yaitu tahap identifikasi konsep dan tahap prediksi relasi. Model AMR *parsing* lintas bahasa memanfaatkan *multilingual word embedding* dalam memahami konteks setiap bahasa yang dibutuhkan. Model stog menggunakan mBERT (Conneau dan Lample, 2019) sebagai *multilingual word embedding*. Terdapat beberapa model lain yang mendukung *multilingual word embedding*, seperti mBART (Liu dkk., 2020) yang merupakan versi *multilingual* dari *Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension* (BART) (Lewis dkk., 2020), dan mT5 (Xue dkk., 2021) yang merupakan versi *multilingual* dari T5 (Raffel dkk., 2020). Teknik AMR *parsing* lintas bahasa dapat memanfaatkan dataset silver yang dapat dibuat dari dataset AMR dan korpus paralel. Korpus paralel pasangan Bahasa Inggris dan Bahasa Indonesia seperti PANL-BPPT (BPPT, 2009) dan IWSLT2017 (Cettolo dkk., 2017) dapat digunakan untuk AMR *parsing* untuk kalimat berbahasa Indonesia.

Putra (2022) menggunakan dataset AMR 2.0 dan korpus paralel PANL-BPPT sebagai data latihnya. Kalimat Bahasa Inggris dari dataset AMR 2.0 ditranslasi menjadi Bahasa Indonesia dan kalimat Bahasa Inggris dari korpus paralel PANL-BPPT dilakukan *parsing* ke graf AMR. Putra (2022) mengevaluasi kualitas dataset dengan menilai kedekatan kalimat hasil translasi dari AMR 2.0 menggunakan *cosine similarity*. Namun, kualitas dataset hasil AMR *parsing* dari korpus paralel tidak dievaluasi. Model terbaik dari penelitian Putra (2022) menghasilkan kinerja SMATCH 51.0. Kinerja tersebut masih kalah dengan *baseline*-nya yang menggunakan teknik *translate-and-parse* dengan kinerja

SMATCH 62.5. Hal ini diduga karena jumlah dan kualitas dataset yang dihasilkan masih belum maksimal, dan model yang dihasilkan belum menggunakan teknik pelatihan model yang lebih baik.

Lee dkk. (2022) memperkenalkan teknik augmentasi data dalam menghasilkan data latih untuk model AMR *parsing* lintas bahasa. Teknik ini menghasilkan lebih banyak pasangan graf AMR dan kalimat dengan bahasa yang dituju, sehingga karena variasi data tersebut, model yang dilatih dapat menghasilkan kinerja yang lebih baik. Teknik tersebut meningkatkan kinerja SMATCH pada model XL-AMR (Blloshmi dkk., 2020) menjadi 63.0 untuk Bahasa Cina, 73.7 untuk Bahasa Jernam, 76.1 untuk Bahasa Italia, dan 77.1 untuk Bahasa Spanyol.

Model AMR *parsing* terbaik saat ini adalah *Graph Pre-training for AMR Parsing and Generation* (AMRBART). AMRBART dibangun dari *language model* BART menggunakan teknik graf *pre-training*. AMRBART merupakan model *state of the art* dengan kinerja SMATCH 85.4 pada dataset AMR 2.0 dan 84.2 pada dataset AMR 3.0 (Bai dkk., 2022).

I.2 Masalah Penelitian

Teknik *cross-lingual* AMR *parsing* terbaru untuk Bahasa Indonesia adalah teknik oleh Putra (2022) yang menggunakan teknik stog oleh (Zhang dkk., 2019). Banyak teknik-teknik lain yang memiliki kinerja AMR *parsing* yang lebih baik, seperti AMRBART (Bai dkk., 2022). Dataset yang digunakan oleh Putra (2022) juga belum menggunakan dataset terbaru, yakni AMR 3.0, yang memiliki kalimat dan aturan anotasi lebih banyak dari AMR 2.0. Terdapat korpus paralel IWSLT2017 (Cettolo dkk., 2017) yang belum digunakan oleh Putra (2022).

Rumusan masalah dari tesis ini adalah bagaimana melakukan AMR *parsing* dari kalimat berbahasa Indonesia menjadi graf AMR berbahasa Inggris dengan menggunakan dataset AMR Bahasa Inggris, korpus paralel Bahasa Inggris-Indonesia, dan *multilingual language model* yang sudah ada.

I.3 Tujuan

Tujuan tesis ini adalah untuk menghasilkan sebuah model AMR *parsing* lintas bahasa dari kalimat berbahasa Indonesia menjadi AMR berbahasa Inggris serta melakukan evaluasi kelayakan model tersebut dalam merepresentasikan makna kalimat berbahasa Indonesia.

I.4 Hipotesis

Premis 1: AMR berbahasa Inggris dapat digunakan sebagai representasi kalimat dalam bahasa lain (Damonte dan Cohen, 2018). Hal ini ditunjukkan dengan analisis kuantitatif berdasarkan nilai SMATCH (Cai dan Knight, 2013) dan kualitatif berdasarkan divergensi translasi (Dorr, 1994).

Premis 2: Teknik graf *pre-training* AMRBART oleh Bai dkk. (2022) yang menggunakan *language model* BART merupakan teknik yang menghasilkan kinerja AMR *parsing* terbaik berdasarkan metrik SMATCH.

Premis 3: Fitur *multilingual word embedding* dan pelatihan pada dataset silver dapat meningkatkan kinerja pembangkit AMR cross-lingual (Blloshmi dkk., 2020) Dataset pasangan kalimat Bahasa Indonesia dan graf AMR berbahasa Inggris dapat dihasilkan dari dataset pasangan teks berbahasa Inggris dan graf AMR yang teks berbahasa Inggrisnya ditranslasi menjadi Bahasa Indonesia serta korpus paralel Indonesia-Inggris yang teks berbahasa Inggrisnya dibangkitkan menjadi graf AMR (Putra, 2022). Teknik augmentasi data oleh Lee dkk. (2022) terbukti dapat meningkatkan kinerja AMR *parsing* lintas bahasa untuk Bahasa Cina, Jerman, Italia, dan Spanyol.

Berdasarkan premis-premis tersebut, disusun hipotesis sebagai berikut:

Hipotesis: Model AMR *parsing cross-lingual* untuk Bahasa Indonesia dapat dibangun dengan menggunakan teknik pre-training AMRBART oleh Bai dkk. (2022) menggunakan *multilingual language model* dengan dataset yang dibangun dari dataset pasangan teks berbahasa Inggris dan graf AMR serta korpus paralel Indonesia-Inggris dengan teknik augmentasi data oleh Lee dkk. (2022). Hasil

kinerja model tersebut dapat menghasilkan kinerja yang lebih baik dari *baseline* yang menggunakan teknik *translate-and-parse* oleh Uhrig dkk. (2021).

I.5 Batasan Masalah

Batasan masalah dari tesis ini adalah:

1. Evaluasi kuantitatif menggunakan metrik SMATCH (Cai dan Knight, 2013) dan kualitatif menggunakan divergensi translasi (Dorr, 1994).

I.6 Metodologi

Berikut adalah tahapan atau metodologi yang dilakukan dalam proses penulisan tesis ini.

1. Analisis Masalah

   Tahap awal dilakukan analisis permasalahan dalam melakukan AMR *parsing* dalam Bahasa Indonesia. Hal ini dilakukan dengan mendalami materi mengenai teknik-teknik AMR *parsing* yang ada sebelumnya dan membandingkan hasil kinerjanya.

2. Perancangan Solusi dan Implementasi

   Pada tahap ini dilakukan perancangan dan implementasi berdasarkan hasil analisis teknik-teknik AMR *parsing* sebelumnya. Dirancang sebuah solusi dengan menggunakan beberapa gabungan teknik dan model tersebut. Lalu dilakukan implementasi dari rancangan tersebut dengan melakukan *training* model terhadap dataset yang tersedia.

3. Pengujian dan Evaluasi

   Hasil implementasi sebelumnya diuji dan dievaluasi secara kuantitatif dengan menggunakan metrik tertentu untuk dibandingkan hasilnya dengan teknik sebelumnya. Hasil AMR *parsing* juga dievaluasi secara kualitatif dengan analisis divergensi translasi. Evaluasi secara kuantitatif dan kualitatif tersebut digunakan untuk kesimpulan kelayakan model AMR *parsing cross-lingual* untuk kalimat berbahasa Indonesia.

Bab II Tinjauan Pustaka

II.1 *Abstract meaning representation* (AMR)

AMR merupakan sebuah bahasa representasi semantik. AMR adalah graf berakar, berlabel, terarah, dan bersiklus yang terdiri dari satu kalimat utuh. Tujuan dari AMR mengabstraksikan kalimat dari representasi sintaktiknya, di mana kalimat-kalimat yang memiliki pengertian yang sama seharusnya memiliki AMR yang sama, walaupun tidak direpresentasikan dengan kalimat yang sama. AMR awalnya dibuat untuk kalimat berbahasa Inggris dan tidak digunakan sebagai bahasa secara umum (Banarescu dkk., 2013).

AMR menggunakan *framesets* PropBank (Kingsbury dan Palmer, 2002) secara ekstensif untuk merepresentasikan predikat yang ada dalam kalimat. Sebagai contoh, kalimat "*the dog likes to eat*" memiliki dua buah predikat di dalamnya, yaitu "*likes*" dan "*eat*". Untuk menggunakan kedua predikat tersebut dalam graf AMR, digunakan frame dari PropBank yang berkorespondensi dengan makna dari kedua predikat dalam kalimat.

Dalam pembentukan graf dari kalimat, AMR tidak mempedulikan urutan dan langkah. Anotasi yang ada pada AMR berisikan aturan-aturan untuk merepresentasikan berbagai macam kata, frasa, dan kalimat. Pada panduan tersebut tidak terdapat aturan eksplisit mengenai langkah pembentukan AMR dari suatu kalimat dan bagaimana urutan yang harus diikuti. Hal ini mendorong peneliti untuk berpikir secara fleksibel mengenai hubungan antara kalimat dan maknanya.

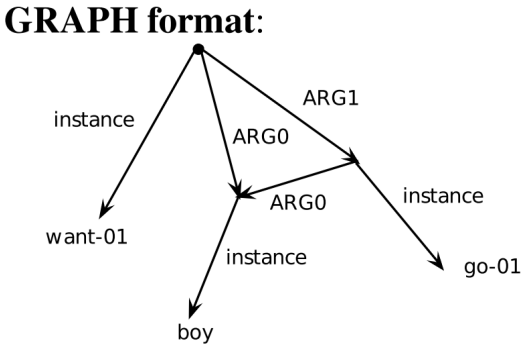
Sebuah AMR dapat dituliskan dalam beberapa format sebagai berikut:

1. Format graf.

   Karena AMR berupa sebuah graf, maka sebuah graf AMR dapat dituliskan dalam sebuah gambar graf. Contoh sebuah graf AMR dari kalimat "*the boy wants to go*" dapat dilihat pada Gambar II.1.

2. Format PENMAN.

   Untuk penulisan secara linear supaya lebih mudah diproses secara komputasi,



Gambar II.1: Contoh sebuah graf AMR dari kalimat "*the boy wants to go*" (Banarescu dkk., 2013).

sebuah AMR dapat dilinearisasi menjadi format PENMAN. Bila suatu entitas digunakan secara berulang, maka digunakan variabel sebagai referensi sebuah simpul pada graf yang dapat digunakan ulang. Contoh sebuah AMR dari kalimat "*the boy wants to go*" dalam format PENMAN dapat dilihat pada Gambar II.2.

AMR format (based on PENMAN):

```
(w / want-01
  :arg0 (b / boy)
  :arg1 (g / go-01
          :arg0 b))
```

Gambar II.2: Contoh sebuah AMR dari kalimat "*the boy wants to go*" dalam format PENMAN (Banarescu dkk., 2013).

3. Format logika.

   Format logika merepresentasikan keterhubungan antar simpul dan/atau sisi dari graf. Contoh sebuah AMR dari kalimat "*the boy wants to go*" dalam format logika dapat dilihat pada Gambar II.3.

LOGIC format:

$\exists$ w, b, g:
instance(w, want-01) $\wedge$ instance(g, go-01)
instance(b, boy) $\wedge$ arg0(w, b) $\wedge$
arg1(w, g) $\wedge$ arg0(g, b)

Gambar II.3: Contoh sebuah AMR dari kalimat "*the boy wants to go*" dalam format logika (Banarescu dkk., 2013).

II.1.1 *Evaluation Metric for Semantic Feature Structures* (SMATCH)

SMATCH merupakan metrics untuk mengukur kesamaan antara dua representasi semantik. Tujuan utama dari semantik parsing adalah untuk menghasilkan hubungan relasi semantik dalam teks. Hasil dari semantik parsing ini biasanya direpresentasikan oleh struktur semantik suatu kalimat utuh. Evaluasi dari struktur ini diperlukan untuk tugas parsing semantik dan tugas anotasi semantik yang menghasilkan linguistic resource untuk semantic parsing. Secara definisi, SMATCH merupakan nilai f-score maksimal yang didapatkan berdasarkan padanan satu-satu antara dua representasi semantik (Cai dan Knight, 2013).

Pada kasus ini, SMATCH digunakan untuk mengukur kesamaan antara dua AMR. Sebagai contoh, terdapat dua kalimat representasi AMR yaitu "*the boy wants to go*" dan "*the boy wants the football*". Perhitungan SMATCH dari kedua AMR ini memerlukan kedua AMR dalam bentuk format logika proposisi yang terdiri dari relasi(variabel, konsep) dan relasi(variabel1, variabel2). Setelah itu dapat dihitung presisi, recall, dan f-score dari kesamaan logika proposisi kedua AMR tersebut.

Sebagai contoh, terdapat dua kalimat "*the boy wants to go*" dan "*the boy wants the football*". Kedua kalimat tersebut diubah menjadi bentuk AMR dengan format logika. Logika proposisi dari representasi AMR pertama dapat dilihat pada Gambar II.4. Sedangkan bentuk logika proposisi untuk representasi kedua dapat dilihat pada Gambar II.5. Berdasarkan contoh tersebut, terdapat 6 cara untuk memasangkan variabel-variabel antara kedua AMR tersebut. Table II.1 menunjukkan kombinasi cara memasangkan variabel kedua AMR tersebut.

```
instance(a, want-01)  ∧
instance(b, boy)  ∧
instance(c, go-01)  ∧
ARG0(a, b)  ∧
ARG1(a, c)  ∧
ARG0(c, b)
```

Gambar II.4: Logika proposisi untuk AMR dari kalimat "*the boy wants to go*".

```
instance(x, want-01)  ∧
instance(y, boy)  ∧
instance(z, football)  ∧
ARG0(x, y)  ∧
ARG1(x, z)
```

Gambar II.5: Logika proposisi untuk AMR dari kalimat "*the boy wants the football*".

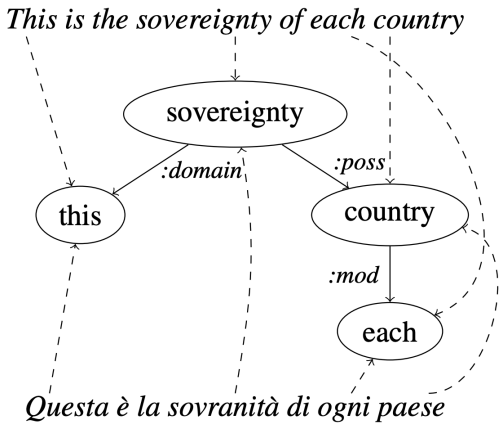
Tabel II.1: Kombinasi pasangan dan kalkulasi F-Score kedua AMR (Cai dan Knight, 2013). Hasil skor SMATCH diambil dari nilai F-Score tertinggi, yaitu 0.73.

Pemetaan	Match	Precision	Recall	F-Score
x=a, y=b, z=c	4	4/5	4/6	0.73
x=a, y=c, z=b	1	1/5	1/6	0.18
x=b, y=a, z=c	0	0/5	0/6	0.00
x=b, y=c, z=a	0	0/5	0/6	0.00
x=c, y=a, z=b	0	0/5	0/6	0.00
x=c, y=b, z=a	2	2/5	2/6	0.36

Karena kedua AMR ini memiliki variabel yang berbeda dan tidak ada informasi mengenai hubungan antar variabel yang ada, perlu dicari semua kemungkinan dari pemetaan variabel yang ada. Dari seluruh pemetaan, nilai SMATCH akan mengambil hasil yang memiliki f-score tertinggi. Untuk mengurangi waktu yang dibutuhkan dalam evaluasi tanpa menurunkan akurasi, dimanfaatkan metode hill-climbing. Metode ini bekerja secara greedy dan tidak seoptimal bruteforce, namun bisa menghasilkan perhitungan yang cukup baik.

II.2 AMR *parsing* Lintas Bahasa

Sebuah AMR merupakan graf dengan node yang merepresentasikan konsep dari kalimat dan edge yang merepresentasikan hubungan semantik antar kalimat. Dataset AMR yang cukup besar untuk di-*train* berisikan pasangan kalimat berbahasa Inggris dengan graf AMR. Properti *cross-lingual* dari AMR di berbagai bahasa merupakan subjek yang sering dibahas. Banarescu dkk. (2013) menyatakan bahwa AMR bukanlah representasi lintas bahasa dan dapat dikategorikan menjadi jenis AMR yang berbeda untuk anotasi dari bahasa yang berbeda. Gambar II.6 merupakan contoh graf AMR dengan pasangan kalimat Bahasa Inggris dan Italia.



Gambar II.6: Contoh graf AMR dengan pasangan kalimat Bahasa Inggris dan Italia (Damonte dan Cohen, 2018).

Tujuan dari AMR adalah untuk mengabstraksikan realisasi sintaktik dari kalimat asli serambi mempertahankan makna yang tersirat. Sebagai konsekuensi, perbedaan frasa yang berbeda dari satu kalimat diharapkan untuk memberikan representasi AMR yang identik. Namun, hal ini tidak selalu berlaku untuk lintas bahasa. Dua kalimat yang kalimat yang mengekspresikan makna yang sama dalam dua yang sama dalam dua bahasa yang berbeda tidak menjamin untuk menghasilkan struktur AMR yang identik (Xue dkk., 2014). Dalam mengatasi permasalahan ini, Damonte dan Cohen (2018) mengusulkan dua metode yang berbeda.

1. Proyeksi anotasi. Setiap simpul pada graf AMR dapat dipasangkan dengan kata pada kalimat berbahasa Inggris. Dengan asumsi bahasa lain memiliki struktur yang mirip dengan Bahasa Inggris, maka dapat dilakukan pemasangan terhadap bahasa lain tersebut.

2. Mesin translasi. Pada metode ini digunakan sebuah mesin translasi untuk menerjemahkan kalimat dengan bahasa lain menjadi Bahasa Inggris. Hasil kalimat berbahasa Inggris tersebut kemudian diubah menjadi graf AMR. Kualitas dari hasil graf AMR tersebut bergantung pada kualitas mesin translasi yang digunakan.

Blloshmi dkk. (2020) mengusulkan konfigurasi dalam melakukan training untuk AMR *parsing cross-lingual*. Konfigurasi tersebut adalah sebagai berikut:

1. Zero-shot. Model dilatih pada kalimat berbahasa Inggris saja dengan mengandalkan fitur *multilingual*, dan dievaluasi pada bahasa yang dituju.
2. Language-specific. Model dilatih pada kalimat berbahasa bahasa yang dituju, misal Bahasa Indonesia, dan dievaluasi pada bahasa yang dituju tersebut.
3. Bilingual. Model dilatih pada kalimat berbahasa Inggris dan bahasa yang dituju. dan dievaluasi pada bahasa yang dituju tersebut.
4. Multilingual. Model dilatih pada kalimat-kalimat dengan berbagai macam bahasa, dan dievaluasi pada bahasa-bahasa yang dituju tersebut.

II.2.1 Divergensi translasi

Divergensi translasi yang diformalisasikan oleh Dorr (1994) digunakan pada penelitian Damonte dan Cohen (2018), Blloshmi dkk. (2020), dan Putra (2022) sebagai dasar dalam melakukan evaluasi secara kualitatif pada hasil AMR *parsing* lintas bahasa. Blloshmi dkk. (2020) mendefinisikan 7 kategori divergensi translasi sebagai berikut:

1. Divergensi *thematic*. Divergensi yang terjadi ketika struktur argumen-predikat yang berbeda pada bahasa yang berbeda. Contoh: "*I like travelling*" dengan kata "*I*" pada Bahasa Inggris merupakan sebuah subjek menjadi "*Mi piace viaggiare*" dengan kata "*Mi*" merupakan sebuah objek pada Bahasa Italia.
2. Divergensi *promotional* dan *demotional*. Divergensi yang terjadi ketika pergantian jenis *modifier* pada bahasa yang berbeda. Contoh: "*John usually goes home*" dengan "*usually*" merupakan sebuah frasa *adverb* menjadi "*Juan suele ir a casa*" dengan "*suele*" merupakan sebuah frasa *verb* pada Bahasa Spanyol.
3. Divergensi *structural*. Divergensi yang terjadi ketika perubahan jenis frasa pada bahasa yang berbeda. Contoh: "*I saw John*" dengan "*John*" merupakan sebuah *noun phrase* (NP) pada Bahasa Inggris menjadi "*Vi a Juan*" dengan "*a Juan*" merupakan sebuah *prepositional phrase* (PP) pada Bahasa Spanyol.

4. Divergensi *conflational*. Divergensi yang terjadi ketika suatu kata menjadi dua atau lebih kata dalam bahasa lain. Contoh: "*I fear*" pada Bahasa Inggris menjadi "*Io ho paura*" (*I have fear*) pada Bahasa Italia.

5. Divergensi *categorical*. Divergensi yang terjadi ketika arti yang sama dituliskan dengan kategori sintaks yang berbeda pada bahasa yang berbeda. Contoh: "*I agree*" dengan "*agree*" merupakan sebuah *verb* pada Bahasa Inggris menjadi "*Estoy de acuerdo*" dengan "*acuerdo*" merupakan sebuah *noun* pada Bahasa Spanyol.

6. Divergensi *lexical*. Divergensi yang terjadi ketika *verb* pada suatu bahasa ditranslasi menjadi *verb* leksikal yang berbeda di bahasa lain. Contoh: "*John broke into the room*" pada Bahasa Inggris menjadi "*Juan forzo la entrada al cuarto*" pada Bahasa Spanyol. Kata "*break*" ditranslasi menjadi "*force*" pada Bahasa Spanyol.

II.3 *Language Model*

Teks yang dapat dipahami oleh manusia tidak dapat dipahami secara langsung oleh model NLP sehingga diperlukan sebuah bentuk representasi lain. Ada beberapa teknik dalam membuat representasi dari sebuah teks, seperti *bag of words*, *n-gram*, *word embedding*, SRL, dan AMR. Dari representasi-representasi tersebut, dapat dilakukan berbagai macam *task* seperti peringkasan teks, klasifikasi sentimen, mesin translasi, *question-answering*, deteksi parafrasa, dan lain-lain.

Word embedding merupakan representasi vektor berdimensi tinggi yang merepresentasikan kedudukan sebuah kata di antara kata-kata lain. *Word embedding* dapat dibentuk dengan berbagai metode. Ada metode yang tidak memberikan konteks pada kata seperti Word2Vec (Mikolov dkk., 2013), GloVe (Pennington dkk., 2014), dan FastText (Bojanowski dkk., 2017). Metode tersebut menghasilkan yang disebut dengan *non-contextual word embedding*. Ada juga metode yang memberikan konteks pada kata dari sebuah kalimat, yang disebut dengan *contextual word embedding*. Dalam memahami konteks sebuah kata dari sebuah kalimat, diperlukan untuk memahami membaca keseluruhan kalimat

tersebut.

Recurrent neural network (RNN), *long short-term memory* (LSTM), dan *gated recurrent neural network* (GRU) kerap digunakan sebagai pendekatan sequence modeling dan permasalahan transduction seperti language modeling dan machine translation. Pendekatan tersebut berguna untuk memahami konteks sebuah kata pada suatu kalimat. Transformer merupakan sebuah sequence-to-sequence model yang terdiri dari *encoder* dan *decoder* (Vaswani dkk., 2017). Transformer bergantung sepenuhnya terhadap mekanisme atensi untuk menggambarkan ketergantungan global antara input dan output. Model berbasis transformer dapat berjalan secara paralel sehingga dapat mempercepat proses training.

Salah satu pengembangan *encoder* dari model transformer adalah *Bidirectional Encoder Representations from Transformers* (BERT). BERT merupakan teknik machine learning berbasis transformer untuk NLP yang dikembangkan oleh Google (Devlin dkk., 2019). BERT memiliki kemampuan untuk memahami konteks dalam sebuah kalimat dan menggunakannya untuk menghasilkan hasil yang lebih akurat daripada model pembelajaran mesin sebelumnya. BERT merupakan pengembangan komponen *encoder* pada model transformer. BERT dilatih terhadap dua *task* yaitu *language modeling* dan *next sentence prediction*. Sebagai hasil dari proses pembelajaran ini adalah, BERT mempelajari embedding kontekstual untuk kata-kata yang ada.

Kemudian dikembangkan BART yang mampu melakukan denoising autoencoder yang memapping dokumen terkorupsi terhadap dokumen originalnya (Lewis dkk., 2020). BART diimplementasikan sebagai model berbasis transformer dengan bidirectional encoder dan left-to-right autoregressive decoder. BART memiliki arsitektur *encoder* seperti BERT (Devlin dkk., 2019) dan *decoder* seperti *Generative Pre-trained Transformer 2* (GPT2) (Radford dkk., 2019). BART (Lewis dkk., 2020) adalah auto-encoder yang diimplementasikan sebagai model berbasis seq2seq pada standar arsitektur Transformer (Vaswani dkk., 2017). BART di-train untuk merekonstruksi teks original berbasis teks terkorupsi yang digenerasi oleh 5 fungsi

noising berikut ini.

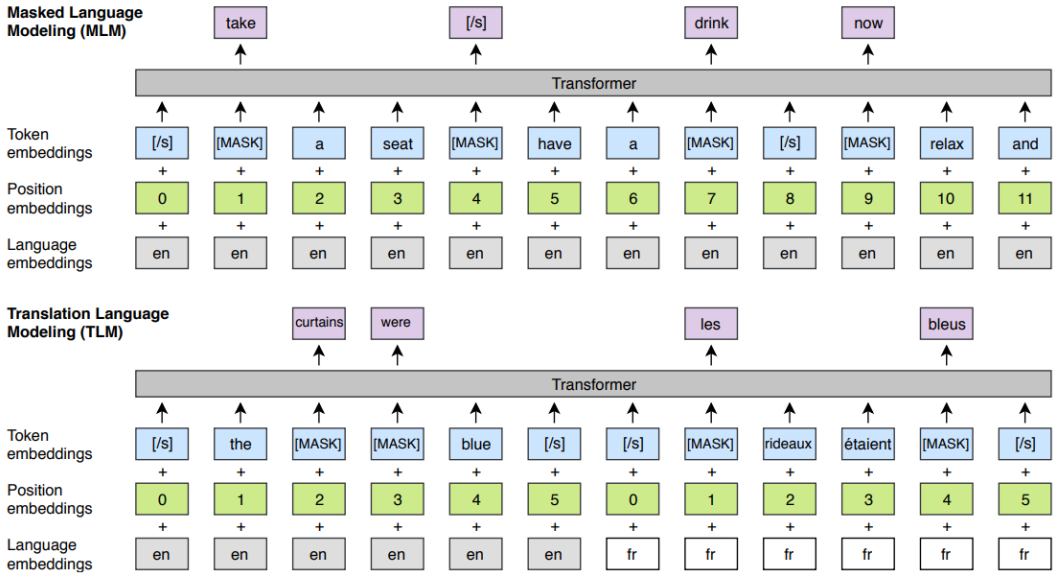
1. Token masking. Token secara random diubah menjadi elemen mask
2. Token deletion. Token secara random dihapus dari input
3. Text Infilling. Teks spans secara random diubah menjadi token single mask
4. Sentence permutation. Teks dibagi menjadi segmen-segmen lalu diacak.
5. Document Rotation. Dokumen dirotasi untuk dimulai dengan token random.

Dalam menangani *task* yang memerlukan lebih dari satu bahasa, seperti mesin translasi, digunakan representasi *multilingual word embedding*. Model *multilingual* dapat dibangun dengan menambahkan *vocabulary* dari bahasa lain dalam data *training*. BERT juga dapat digunakan untuk menghasilkan *multilingual word embedding* dengan dilatih menggunakan dataset yang terdiri dari 101 bahasa untuk menghasilkan *multilingual* model bernama mBERT. Terdapat variasi BERT yang digunakan sebagai *multilingual word embedding* yang tidak terikat ke bahasa apapun, yaitu *Language-agnostic BERT Sentence Embedding* (LABSE) (Feng dkk., 2022). LABSE berusaha merepresentasikan kalimat dalam bahasa yang berbeda-beda menjadi satu representasi yang sama. BART juga dikembangkan untuk menghasilkan representasi *multilingual word embedding* bernama mBART (Liu dkk., 2020).

Model berbasis transformer lain bernama T5 (Raffel dkk., 2020) merupakan model yang dilatih dengan campuran *task unsupervised* dan *supervised*. T5 bekerja dengan baik untuk berbagai macam *task* seperti peringkasan teks dan translasi. Model T5 juga dikembangkan untuk *multilingual word embedding* bernama mT5 (Xue dkk., 2021) yang dilatih pada Common Crawl yang memiliki 101 bahasa.

Cross-lingual Language Model (XLM) menunjukkan sebuah teknik *pre-training* lain untuk memanfaatkan korpus paralel. XLM memanfaatkan kombinasi *Masked Language Modeling* (MLM) dan *Translation Language Modeling* (TLM). Teknik *pre-training* MLM mengikuti cara *pre-training* BERT, yaitu dengan melakukan *masking* pada beberapa kata dari input. Teknik *pre-training* TLM melakukan

concatenation antara dua kalimat dari bahasa yang berbeda. Kalimat dari bahasa yang berbeda tersebut dianggap sebagai *next sentence* pada model BERT dengan mengulang kembali *position embeddings* dari indeks awal kembali. Ilustrasi *pre-training* XLM dapat dilihat pada Gambar II.7.



Gambar II.7: *Cross-lingual Language Model* (XLM) *pre-training* (Conneau dan Lample, 2019).

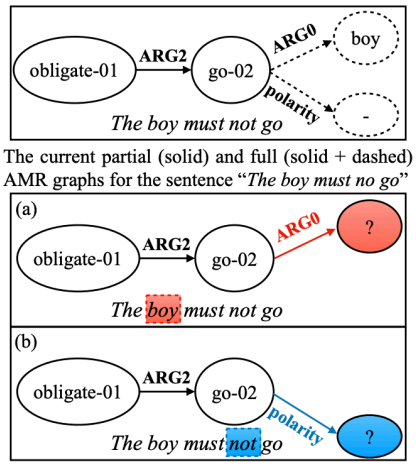
II.4 Penelitian Terkait

II.4.1 AMR Parsing via Graph-sequence Iterative Inference (Cai dan Lam, 2020)

Salah satu tantangan dalam AMR *parsing* adalah kurangnya pemetaan eksplisit antara simpul pada graf dan kata kata dalam teks. Untuk saat ini, akurasi *parsing* dari penelitian-penelitian terkait masih belum memuaskan dibandingkan kinerja manusia, terutama pada kasus dimana kalimat lebih panjang dan informatif. Salah satu kemungkinan alasan dari kekurangan ini adalah kurangnya model interaksi antara prediksi konsep dan prediksi relasi yang penting untuk mendapatkan hasil yang tidak ambigu.

Pada tingkat dasar, kita dapat mengkategorikan pendekatan AMR parsing menjadi dua kelas, yaitu dua tahap *parsing* dan satu tahap *parsing* (Cai dan Lam, 2020). Pada pendekatan dua tahap *parsing* digunakan desain *pipeline* untuk identifikasi

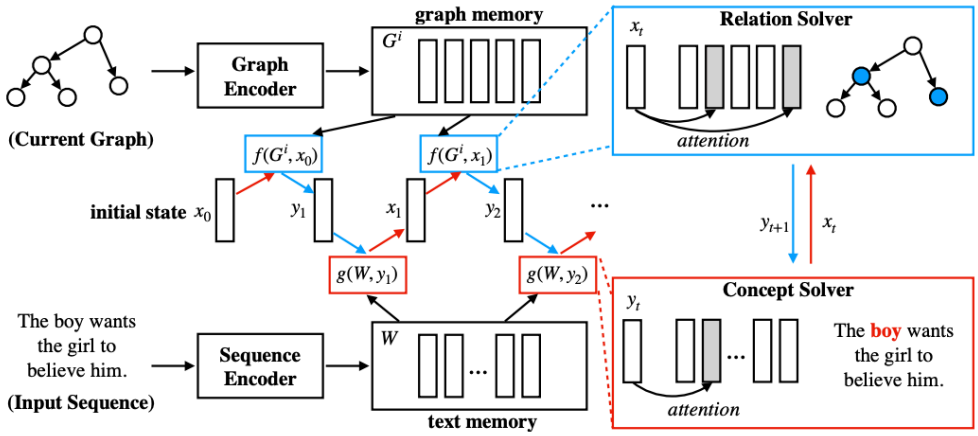
keseluruhan konsep, lalu diikuti dengan prediksi relasi antar hasil prediksi konsep. Pada pendekatan satu tahap *parsing* dikategorikan menjadi tiga jenis, yaitu *parsing* berbasis transisi, *parsing* berbasis seq2seq, dan *parsing* berbasis graf. Pendekatan satu tahap jenis *parsing* berbasis transisi dilakukan dengan memproses kalimat dari kiri ke kanan dan membangun grafik secara bertahap dengan secara bergantian memasukkan simpul atau sisi baru. Pendekatan satu tahap jenis *parsing* berbasis seq2seq dengan melihat *parsing* sebagai transduksi urutan linear ke urutan linear juga dengan memanfaatkan linearisasi grafik AMR. Pendekatan satu tahap jenis *parsing* berbasis graf di mana setiap langkah waktu, simpul baru beserta koneksinya ke simpul yang ada diputuskan bersama secara berurutan maupun paralel. Gambar II.8 merupakan contoh proses pembentukan graf AMR dengan pendekatan satu tahap jenis *parsing*.



The current partial (solid) and full (solid + dashed) AMR graphs for the sentence "*The boy must no go*"

Gambar II.8: Pembangunan sebuah graf AMR dari subgraf AMR yang sebagian terbentuk (Cai dan Lam, 2020). Lanjutan kemungkinan tahap pembangunan dapat berupa ekspansi untuk: (a) konsep "*boy*" dengan relasi ARG0 atau (b) konsep negasi dengan relasi polarity.

Cai dan Lam (2020) mengusulkan pendekatan AMR *parsing* via *graph-sequence iterative inference* yang meniru proses manusia dalam melakukan deduksi graf semantik dari suatu kalimat. Pendekatan ini dimulai dari graf kosong yang memanjang secara iteratif dari simpul ke simpul. Ilustrasi pendekatan ini dapat dilihat pada Gambar II.9.



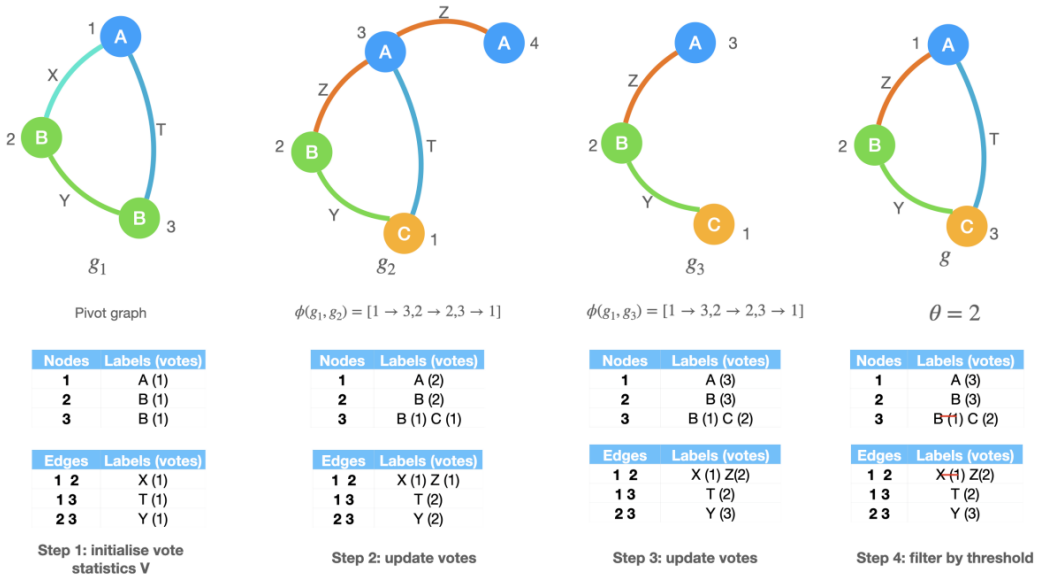
Gambar II.9: Ilustrasi dari pendekatan *graph-sequence iterative inference* untuk AMR *parsing* (Cai dan Lam, 2020).

II.4.2 Teknik *ensemble* untuk AMR *parsing*

Kenaikan kinerja dari AMR *parsing* dari penelitian-penelitian sebelumnya tidak lagi meningkat secara signifikan (Lee dkk., 2022). Ini dikarenakan efek dari self-learning dan augmentasi silver data mulai berkurang. Untuk mengatasi hal tersebut, Lee dkk. (2022) mengusulkan untuk menggabungkan teknik *ensemble* berbasis SMATCH (Hoang dkk., 2021) dengan *ensemble distillation* untuk menghasilkan silver data berkualitas tinggi.

Teknik *ensemble* yang diusulkan oleh Hoang dkk. (2021) bernama Algoritma Graphene dapat dilihat pada Gambar II.10. Ide utamanya adalah dengan memperbaiki sebuah pivot graf berdasarkan graf yang dihasilkan oleh *parser* lain. Dari pivot graf, dilakukan *voting* untuk menghitung jumlah simpul dan sisi graf yang berkorelasi. Setelah dilakukan *voting*, dipilih simpul dan sisi yang memiliki jumlah voting terbesar sebagai hasil akhir graf tersebut. Hal ini dilakukan berulang kali untuk masing-masing graf sebagai pivot awal. Tahapan algoritma ini dapat dilihat sebagai pseudo-code pada Algorithm 1

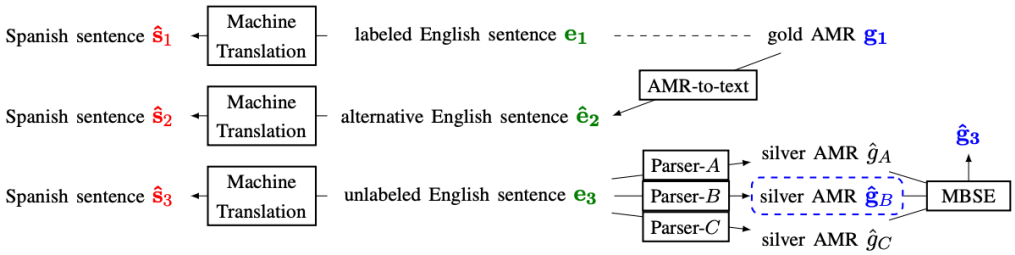
Lee dkk. (2022) mengusulkan sebuah framework untuk melakukan augmentasi data AMR. Ilustrasi framework tersebut dapat dilihat pada Gambar II.11. Dengan keterbatasan jumlah dataset yang dipunya untuk melatih model *cross-lingual*, framework ini menghasilkan lebih banyak pasangan graf AMR dan kalimat dengan



Gambar II.10: Ilustrasi teknik *ensemble* (Hoang dkk., 2021).

bahasa yang dituju. Teknik-teknik yang digunakan untuk menghasilkannya adalah sebagai berikut:

1. Kalimat berbahasa Inggris dari dataset AMR ditranslasi menjadi bahasa yang dituju.
2. Graf AMR dari dataset AMR dilakukan *parsing* menjadi kalimat Bahasa Inggris untuk mendapatkan sebuah alternatif kalimat berbahasa Inggris. Dari kalimat alternatif tersebut dilakukan translasi menjadi bahasa yang dituju.
3. Kalimat-kalimat dari dataset kumpulan kalimat berbahasa Inggris tak berlabel ditranslasi menjadi bahasa yang dituju dan dilakukan *parsing* menjadi graf AMR menggunakan teknik *ensemble*.



Gambar II.11: Framework untuk melakukan augmentasi pasangan data AMR dan kalimat berbahasa Inggris untuk menghasilkan silver data (Lee dkk., 2022).

Algorithm 1 Algoritma Graphene untuk *ensemble* graf (Hoang dkk., 2021).

Input: a set of graphs $G = \{g_1, g_2, ..., g_m\}$ and the support threshold θ
Output: an ensemble graph g^e
Algorithm: Graphene(G,θ)
for $i \leftarrow 1$ **to** m **do**
  $g_{pivot} \leftarrow g_i$
  $V \leftarrow$ Initialise(g_{pivot})
  for $j \leftarrow 1$ **to** m **do**
    if $j \neq i$ **then**
      $V \leftarrow V \cup$ getVote($\phi(g_{pivot}, g_j)$)
    end if
  end for
  $g_i^e \leftarrow Filter(V, \theta)$
end for
$g^e \leftarrow$ the graph with the largest support among $g_1^e, ..., g_m^e$
return g^e

II.4.3 *Graph Pre-training for AMR Parsing and Generation* (AMRBART) (Bai dkk., 2022)

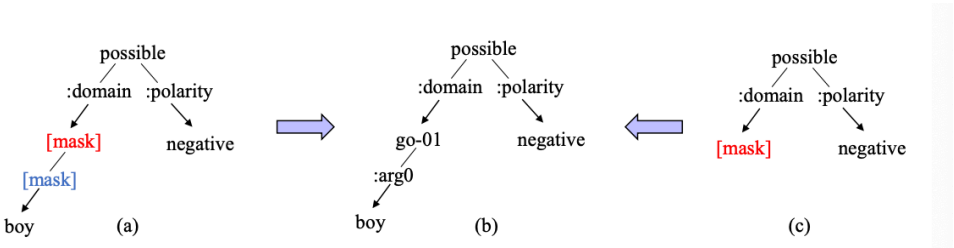
Pretrained language model (PLM) telah terbukti dapat melakukan *task* AMR *parsing* dan generasi AMR-to-text dengan baik. Namun, PLM umumnya dilatih pada data tekstual, sehingga tidak optimal untuk melakukan generasi data terstruktur seperti AMR. Bai dkk. (2022) memperbaiki permasalahan tersebut dengan menambahkan strategi *pre-training* pada model untuk mengintegrasikan informasi teks dan graf AMR. Model ini melinearisasi graf AMR ke sekuens sehingga baik AMR parsing dan AMR-to-text generation dapat dilakukan menggunakan model seq2seq. Model ini melakukan pre-training pada struktur AMR menggunakan BART.

Tahapan *preprocessing* dalam model ini mengadopsi penelitian Bevilacqua dkk. (2021), yaitu *recategorization* untuk mengurangi ukuran *vocab* untuk menangani *sparsity* data. *Recategorization* dilakukan dengan menghapus simpul *sense*, *link* wiki, atribut polaritas, dan menganonimasi *named entity*. Walaupun *recategorization* menunjukkan kinerja yang kurang pada penelitian Bevilacqua dkk. (2021), namun pada AMRBART tidak terjadi pengurangan kinerja. Lemmatisasi juga dilakukan pada fase *preprocessing*. Linearisasi graf menggunakan algoritma

depth-first search (DFS) yang diadopsi dari penelitian Bevilacqua dkk. (2021). Pembentukan graf AMR hasil prediksi dilakukan pada tahapan *postprocessing* dengan mengubah bentuk format DFS menjadi bentuk format PENMAN. Wikifikasi, pemberian *link* pada atribut wiki, juga dilakukan menggunakan BLINK Entity Linker (Wu dkk., 2020). Pengembalian atribut polaritas dilakukan dengan mengidentifikasi lemma yang bersifat negasi.

Model ini mengenalkan dua strategi *self-supervised training* dalam melakukan *pre-training* model BART pada graf AMR. Dapat dilihat pada Gambar II.12, strategi level *denoising* simpul/sisi mendukung model untuk menangkap pengetahuan lokal mengenai simpul dan sisi. Strategi *denoising* level graf mengarahkan model untuk memprediksi sub-graf yang dapat memfasilitasi pembelajaran graf.

1. *Denoising* level simpul/sisi. Pengaplikasian fungsi noise pada simpul dan sisi AMR untuk mengkonstruksi input graf yang kotor. Fungsi noise ini diimplementasikan dengan *masking* 15% simpul dan 15% sisi di setiap graf.
2. *Denoising* level sub-graf. *Task* ini bertujuan untuk mengembalikan graf lengkap ketika diberikan sebagian dari graf. Metode ini menghilangkan sub-graf secara acak dari graf dan mengubahnya dengan token *mask*.



Gambar II.12: Ilustrasi strategi pre-training: (1) *denoising* level simpul/sisi (a->b); (2) *denoising* level sub-graf (c->b) (Bai dkk., 2022).

Strategi *pre-training* (P.T.) dan *fine-tuning* (F.T.) dapat dilihat pada Table II.2. Token <s> dan <g> digunakan untuk membedakan informasi teks dan graf pada input. *Pre-training* dilakukan dengan menghapus sebagian dari input teks/graf dengan teknik *denoising*. *Fine-tuning* dilakukan dengan menghapus keseluruhan dari teks/graf yang dituju. Pada standard *pre-training* dan *fine-tuning*, input berupa salah satu dari teks atau graf dan outputnya berupa teks atau graf yang dituju.

Namun, pada unified *pre-training* dan *fine-tuning*, input berupa kedua teks dan graf, dan outputnya berupa teks atau graf yang dituju.

Tabel II.2: Strategi *pre-training* (P.T.) dan *fine-tuning* (F.T.) untuk pelatihan graf (Bai dkk., 2022). t/g merupakan teks/graf *original*. $\hat{t}/\hat{g}$ merupakan teks/graf yang *noisy* (hasil *denoising*). $\bar{t}/\bar{g}$ merupakan teks/graf yang hilang.

Phase	Task	Input	Output
Std. P.T.	$\hat{t}2t$	<s>x1,..[mask]..,xn</s>	<s>x1,x2,...,xn</s>
	$\hat{g}2g$	<g>g1,..[mask]..,gm</g>	<g>g1,g2,...,gm</g>
Std. F.T.	$g2t$	<g>g1,g2,...,gm</g>	<s>x1,x2,...,xn</s>
	$t2g$	<s>x1,x2,...,xn</s>	<g>g1,g2,...,gm</g>
Unified P.T.	$\hat{t}\bar{g}2t$	<s>x1,..[mask]..,xn</s><g>[mask]</g>	<s>x1,x2,...,xn</s>
	$\bar{t}\hat{g}2g$	<s>[mask]</s><g>g1,..[mask]..,gm</g>	<g>g1,g2,...,gm</g>
	$\hat{t}g2t$	<s>x1,..[mask]..,xn</s><g>g1,g2,...,gm</g>	<s>x1,x2,...,xn</s>
	$t\hat{g}2g$	<s>x1,x2,...,xn</s><g>g1,..[mask]..,gm</g>	<g>g1,g2,...,gm</g>
	$\hat{t}\hat{g}2t$	<s>x1,..[mask]..,xn</s><g>g1,..[mask]..,gm</g>	<s>x1,x2,...,xn</s>
	$\hat{t}\hat{g}2g$	<s>x1,..[mask]..,xn</s><g>g1,..[mask]..,gm</g>	<g>g1,g2,...,gm</g>
Unified F.T.	$\bar{t}g2t$	<s>[mask]</s><g>g1,g2,...,gm</g>	<s>x1,x2,...,xn</s>
	$t\bar{g}2g$	<s>x1,x2,...,xn</s><g>[mask]</g>	<g>g1,g2,...,gm</g>

II.4.4 Peringkasan Abstraktif Multidokumen Menggunakan Abstract Meaning Representation untuk Bahasa Indonesia (Severina dan Khodra, 2019)

Penelitian oleh (Severina dan Khodra, 2019) membahas mengenai pembangunan sistem peringkasan secara abstraktif dari dokumen berbahasa Indonesia dengan memanfaatkan AMR. Pendekatan yang digunakan adalah AMR *parsing* yang memanfaatkan aturan dan kamus untuk membangun graf AMR berbahasa Indonesia dari teks berbahasa Indonesia. Aturan pembangunan graf tersebut adalah sebagai berikut:

1. Simpul akar graf AMR yang dihasilkan ditentukan dengan *dependency parser*.
2. ARG0 dan ARG1 dari suatu predikat ditentukan dengan melihat apakah predikat merupakan kata kerja aktif atau pasif. Argumen akan diisi dengan subjek dan objek dari kata kerja terkait.
3. Keterangan dari predikat akan menjadi ARG2.

Teknik AMR *parsing* ini masih terbatas dengan banyak jumlah relasi ARG hanya sebanyak 3 buah. Konsep pada AMR yang dihasilkan masih berupa predikat yang diberi label "*-01*" di belakangnya dikarenakan PropBank berbahasa Indonesia masih belum ada. Teknik evaluasi yang digunakan pada penilitian ini tidak menggunakan SMATCH, namun menggunakan ukuran akurasi kemunculan simpul pada AMR yang dihasilkan dibandingkan dengan referensinya.

II.4.5 Pembangkitan Graf Abstract Meaning Representation Berbahasa Indonesia (Ilmy dan Khodra, 2020)

Pada penelitian ini, dilakukan pembangkitan graf AMR berbahasa Indonesia menggunakan pembelajaran mesin. Pendekatan yang digunakan adalah pendekatan menggunakan hasil *dependency parsing* untuk mengetahui hubungan antar kata. Dataset yang digunakan untuk melatih pembangkitan graf AMR masih terbatas dan hanya terdiri dari kalimat sederhana. Metode dalam penelitian ini dikembangkan berdasarkan pada penelitian Zhang dkk. (2019). Metode tersebut melibatkan tiga tahap, yaitu:

1. Prediksi pasangan yang menghasilkan kandidat pasangan dengan konsep yang memiliki relasi dari input.
2. Prediksi label yang melakukan prediksi mengenai relasi kandidat pasangan.
3. Melakukan *post-process* dari kandidat pasangan berlabel untuk menghasilkan graf AMR yang valid.

Hasil dari metode ini dapat menghasilkan hasil yang baik untuk kalimat terstruktur yang simpel dengan skor SMATCH 0.820 pada test dataset kalimat sederhana. Namun, masih untuk kalimat yang lebih kompleks, hasil masih belum memuaskan dengan skor SMATCH 0.684 untuk topik b-salah-darat, 0.583 untuk topik c-gedung-roboh, 0.677 topik d-indo-fuji, 0.687 topik untuk fbunuh-diri, and 0.672 untuk topik g-gempa-dieng.

II.4.6 Pembangkitan Abstract Meaning Representation Lintas Bahasa dari Kalimat Berbahasa Indonesia (Putra, 2022)

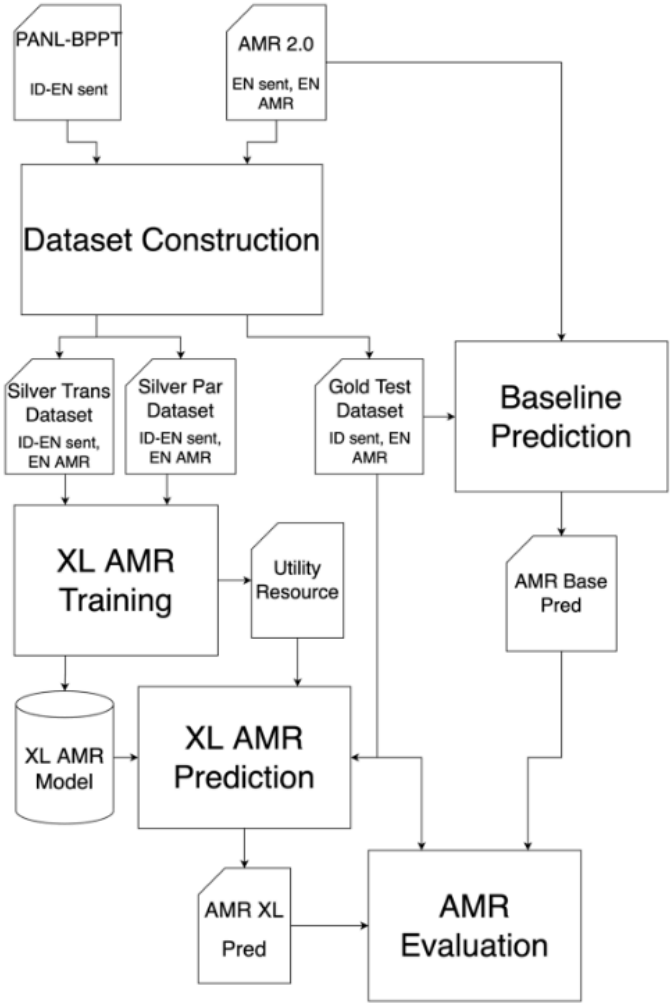
Penelitian ini melakukan AMR *parsing* dari kalimat berbahasa Indonesia dengan pendekatan cross lingual menggunakan dataset berkualitas silver dan gold. Pada pembangkitan ini, dilakukan pemilihan kalimat berdasarkan kedekatan semantiknya menggunakan *cosine similarity* sehingga kalimat dengan kinerja rendah dapat dikeluarkan terlebih dahulu dan mengurangi gap kinerja model. Sumber daya yang digunakan dalam penelitian ini adalah *multilingual word embedding* Numberbatch yang memiliki word embedding untuk bahasa Indo-Malay dan Inggris. Untuk meningkatkan efisiensi dari *word embedding* tersebut, beberapa karakter Cina dan Arab yang tidak muncul dalam dataset AMR 2.0 dihapus. Untuk sumber daya contextual *multilingual word embedding*, digunakan beberapa alternatif berupa mBERT (Devlin dkk., 2019), XLM-R (Conneau dan Lample, 2019), dan mT5 (Xue dkk., 2021).

Gambaran keseluruhan sistem rancangan solusi dapat dilihat pada Gambar II.13. Tahapan yang dilakukan adalah sebagai berikut:

1. Kontruksi korpus sebagai dataset yang dibutuhkan.
2. Pelatihan model AMR *parsing* lintas bahasa.
3. Inferensi kalimat berbahasa Indonesia menjadi graf AMR berbahasa Inggris.

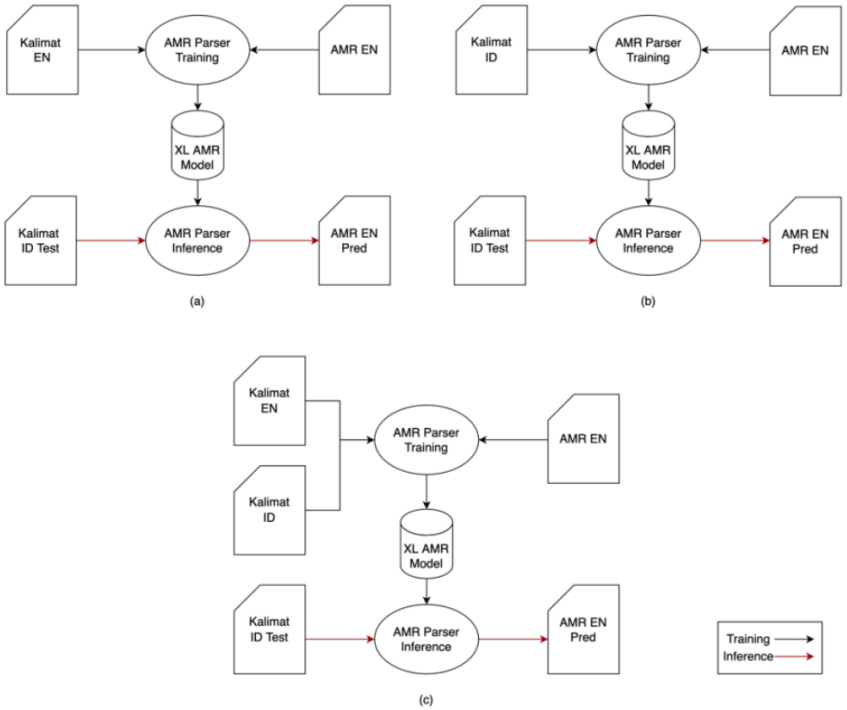
Skema model pelatihan *cross-lingual* AMR *parsing* menggunakan skema yang diusulkan oleh Blloshmi dkk. (2020), yaitu *zero-shot*, *language-specific*, dan *bilingual*. Gambar II.14 menunjukkan tiga skema pelatihan tersebut. *Zero-shot* hanya menggunakan kalimat berbahasa Inggris saja sebagai input model dan mengandalkan *multilingual word embedding* untuk memahami konteks Bahasa Indonesia. *Language-specific* menggunakan kalimat berbahasa Indonesia saja sebagai input model. *Bilingual* menggunakan kalimat berbahasa Inggris dan Indonesia sebagai input model yang kemudian dievaluasi ke Bahasa Indonesia.

Proses evaluasi pada penelitian ini dilakukan dengan pelaksanaan inferensi pada



Gambar II.13: Gambaran keseluruhan sistem rancangan solusi pembangunan model *cross-lingual* untuk Bahasa Indonesia (Putra, 2022).

kalimat berbahasa Indonesia dari dataset gold untuk menghasilkan AMR berbahasa Inggris. Dari graf AMR ini akan dilakukan perhitungan skor SMATCH dengan membandingkannya dengan graf AMR gold. Konfigurasi model AMR *parsing* lintas bahasa yang memiliki skor SMATCH terbaik adalah dengan *multilingual word embedding* dari PLM mT5 dengan skor SMATCH 51.0. Namun teknik *baseline translate-and-parse* masih memiliki skor SMATCH yang lebih tinggi, yaitu 62.5.



Gambar II.14: Model pelatihan *cross-lingual* untuk Bahasa Indonesia: (a) *zero-shot*, (b)*language-specific*, dan (c) *bilingual* (Putra, 2022).

Bab III Analisis Masalah dan Perancangan Solusi

III.1 Analisis Masalah

Pelatihan pembelajaran mesin untuk *task* AMR *parsing* memerlukan dataset *training* pasangan kalimat dan graf AMR-nya. Penelitian Putra (2022) menggunakan dataset silver yang dihasilkan dari AMR 2.0 dan korpus paralel PANL-BPPT sebagai data latih AMR *parsing* lintas bahasa untuk Bahasa Indonesia. Berdasarkan pada penelitian Lee dkk. (2022), model AMR *parsing* dapat meningkat kinerjanya dengan bertambahnya jumlah dataset silver. Dataset yang digunakan Putra (2022) secara kuantitas masih kurang karena ada korpus paralel IWSLT17 dan dataset AMR 3.0 yang belum digunakan.

Pembangunan dataset silver pada penelitian Putra (2022) tidak dievaluasi kualitasnya. Banyak potensi *instance* data berkualitas buruk yang digunakan sebagai data latih, sehingga dapat mengurangi kinerja model. Perhitungan kualitas tersebut juga dapat dimanfaatkan sebagai acuan dalam memperbaiki kualitas dataset silver untuk meningkatkan kinerja model AMR *parsing*.

Model AMRBART (Bai dkk., 2022) merupakan model *state of the art* yang dibangun untuk AMR *parsing* dan AMR-to-text untuk Bahasa Inggris. Belum dikembangkan sebuah model AMR *parsing* untuk lintas bahasa berdasarkan model graf pralatih AMRBART.

III.2 Analisis Solusi

Terdapat dua aspek dari permasalah dataset yang dapat diperbaiki dari penelitian Putra (2022), yaitu jumlah dan kualitas dataset. Pada AMR *parsing* lintas bahasa, dataset silver diperlukan untuk membangun model AMR *parsing* lintas bahasa yang bisa didapatkan dari dataset gold yang telah ada dalam Bahasa Inggris. Putra (2022) telah membangun dataset silver untuk Bahasa Indonesia dari dataset AMR 2.0 dan korpus paralel BPPT-PANL (BPPT, 2009). Jumlah dataset silver dapat diperbanyak dengan cara-cara berikut:

1. Menggunakan dataset AMR 3.0. Dataset AMR 3.0 memiliki sekitar 20,000

pasangan kalimat dan graf AMR lebih banyak dibandingkan dengan Dataset AMR 2.0.

2. Menggunakan korpus paralel tambahan IWSLT17. Putra (2022) hanya menggunakan korpus paralel BPPT-PANL (BPPT, 2009). Korpus paralel IWSLT17 (Cettolo dkk., 2017) dapat menambah sampai 107,329 pasang dataset silver.

3. Mengadopsi teknik augmentasi data oleh Lee dkk. (2022). Melakukan *parsing* graf AMR dari dataset gold menjadi teks berbahasa Inggris, lalu dilakukan translasi ke Bahasa Indonesia. Teknik ini dapat menambah jumlah pasangan dataset silver dari dataset AMR menjadi sampai maksimal dua kali lipat lebih banyak.

Peningkatan kualitas dataset silver dapat dilakukan dengan menggunakan teknik *ensemble* oleh Hoang dkk. (2021). Berdasarkan hasil penelitian Lee dkk. (2022), teknik *ensembling* 5 model AMR *parser* dapat meningkatkan kualitas dataset silver secara signifikan.

Kualitas dataset silver dapat diukur dengan menghitung *cosine similarity* dari *multilingual sentence embedding* dari pasangan kalimat Bahasa Indonesia dan Bahasa Inggris (Blloshmi dkk., 2020). Model *multilingual word embedding* yang digunakan untuk mendapatkan *multilingual sentence embedding* dari sebuah kalimat adalah LABSE (Feng dkk., 2022). Dalam meningkatkan kualitas dataset silver tersebut, dapat dilakukan dilakukan filtrasi untuk menghapus hasil translasi yang buruk. Filtrasi dilakukan dengan mengaplikasikan algoritma *1-Nearest Neighbour* (1-NN) dengan *cosine similarity* sebagai nilai kedekatan antar kalimat berbahasa Inggris dan hasil translasinya. Apabila hasil kalimat yang ditemukan dari 1-NN bukan merupakan kalimat awal sebelum translasi, maka pasangan kalimat tersebut tidak digunakan sebagai data pelatihan.

Teknik *pre-training* graf oleh Bai dkk. (2022) yang menggunakan *language model* BART terbukti dapat meningkatkan skor SMATCH untuk AMR *parsing*. Teknik tersebut belum dicoba untuk *cross-lingual* AMR *parsing*. Skema model

pelatihan *cross-lingual* AMR *parsing* telah diusulkan oleh Blloshmi dkk. (2020), yaitu *zero-shot*, *language-specific*, dan *bilingual*. Pada penelitian Putra (2022), ditunjukkan bahwa skema *bilingual* menghasilkan kinerja terbaik. Teknik *pre-training* graf oleh Bai dkk. (2022) tersebut dapat dilakukan dengan skema pelatihan *bilingual* (Blloshmi dkk., 2020) untuk membangun model *cross-lingual* untuk *task* AMR *parsing*.

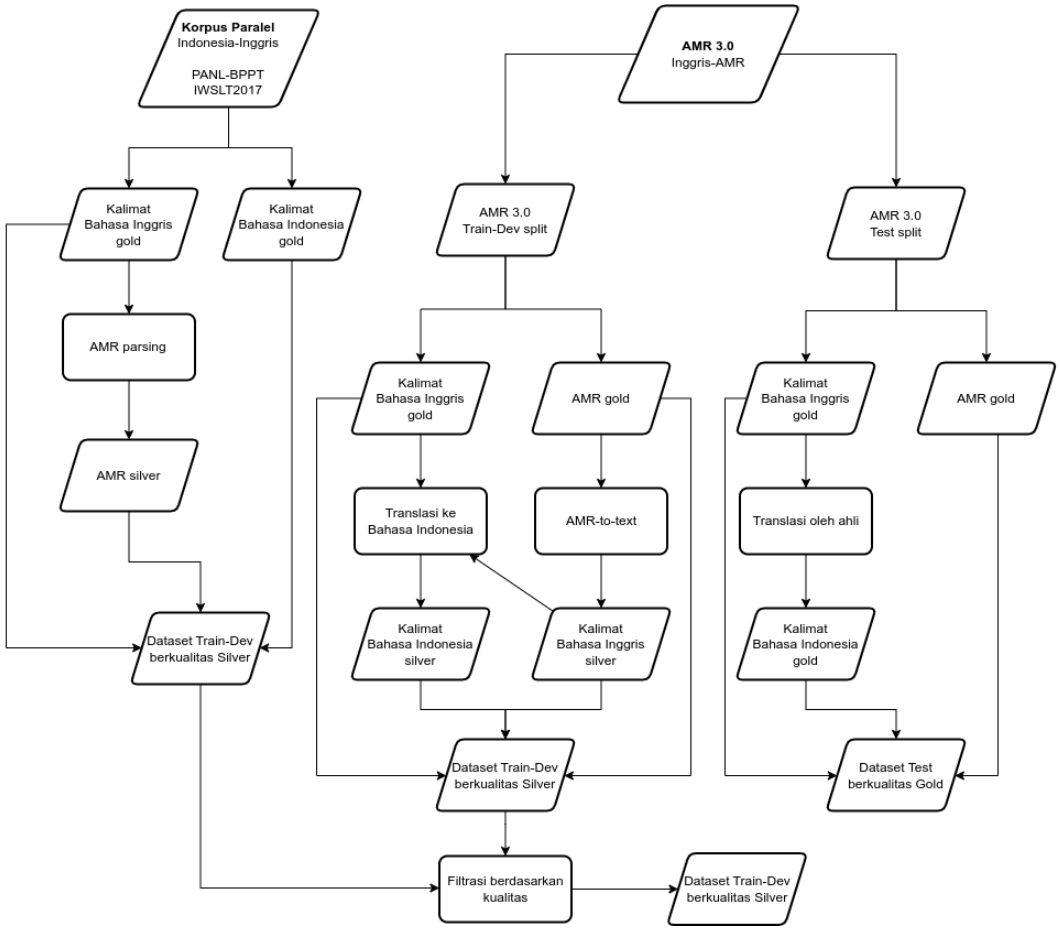
Skema *bilingual* pada penelitian Blloshmi dkk. (2020) dan Putra (2022), dataset silver berbahasa Inggris dan Indonesia dianggap sebagai *instance* data yang berbeda. Mengadopsi teknik *pre-training* XLM, kalimat berbahasa Inggris dan Indonesia dapat di-*concatenate* menjadi satu *instance* data.

III.3 Rancangan Solusi

Pada masalah kurangnya jumlah dataset yang digunakan sebagai data latih model, dibangun kumpulan dataset silver untuk pelatihan model AMR *parsing* lintas bahasa. Diagram alur keseluruhan proses konstruksi dataset gold dan silver dapat dilihat pada Gambar III.1. Alur proses konstruksi dataset berkualitas silver mengikuti framework augmentasi data oleh Lee dkk. (2022).

Korpus paralel PANL-BPPT dan IWSLT2017 digunakan sebagai dataset pasangan kalimat Bahasa Inggris dan Bahasa Indonesia. Kalimat berbahasa Inggris dari korpus paralel diubah menjadi graf AMR. Model AMR *parsing* yang digunakan adalah teknik *ensembling* dari 5 model *state of the art* (Lee dkk., 2022). Hasil graf AMR yang dihasilkan akan digunakan sebagai dataset latih berkualitas silver.

Dataset AMR berkualitas gold yang digunakan adalah Dataset AMR 3.0 (LDC2020T02). Augmentasi data untuk konstruksi dataset berkualitas silver mengambil bagian Train dan Dev *split* dari AMR 3.0. Kalimat berbahasa Inggris dari dataset AMR tersebut diubah menjadi kalimat berbahasa Indonesia dengan menggunakan mesin translasi. Model mesin translasi yang digunakan adalah model OPUS-MT (Tiedemann dan Thottingal, 2020). Graf AMR berkualitas gold juga dilakukan *parsing* menjadi kalimat berbahasa Inggris (AMR-to-text) untuk menambah jumlah variasi pasangan dataset. Model AMR-to-text yang digunakan



Gambar III.1: Diagram alur untuk konstruksi dataset gold dan silver dari korpus paralel dan dataset AMR.

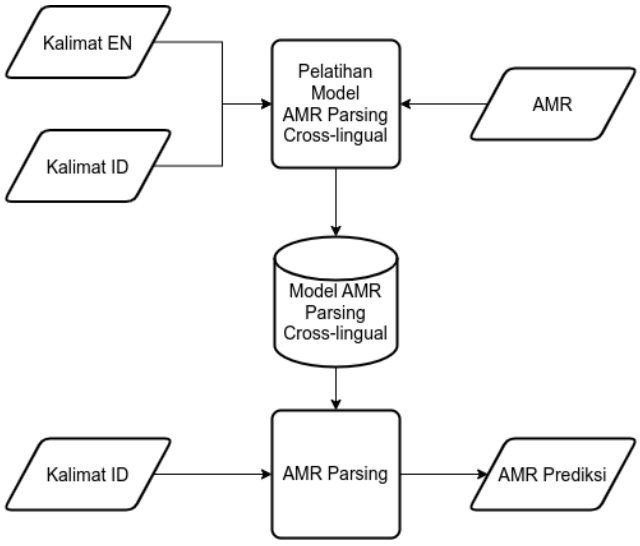
adalah AMRBART. Kalimat berbahasa Inggris hasil *parsing* tersebut juga diubah menjadi kalimat berbahasa Indonesia dengan menggunakan mesin translasi. Maka akan terdapat dua pasangan kalimat berbahasa Inggris dan Indonesia untuk satu buah graf AMR. Dua pasangan tersebut dilakukan *cross join* untuk menghasilkan empat pasang data latih berkualitas silver.

Kumpulan semua dataset silver tersebut kemudian dilakukan filtrasi dengan algoritma 1-NN dengan *cosine similarity* sebagai fungsi jarak antar representasi kalimatnya. Hasil filtrasi yang dihasilkan akan digunakan sebagai dataset latih. Dataset latih tersebut juga dihitung kualitasnya menggunakan rata-rata *cosine similarity*.

Augmentasi data untuk konstruksi dataset berkualitas gold mengambil bagian Test *split* dari AMR 3.0. Bagian kalimat berbahasa Inggris dari Test *split* ditranslasi

menjadi Bahasa Indonesia oleh ahli. Hasil kalimat berbahasa Indonesia tersebut berkualitas gold dan digunakan sebagai dataset test.

Skema pelatihan yang digunakan hanya untuk yang melakukan AMR *parsing* dengan menggunakan skema *bilingual*. Ilustrasi pelatihan dengan skema *bilingual* dapat dilihat pada Gambar III.2 Pelatihan dicoba dengan konfigurasi dengan dan tanpa *concatenating*. Pada konfigurasi tanpa *concatenating* tetap mengikuti teknik pelatihan sebelumnya yang dapat dilihat pada Table III.1. Sedangkan pada konfigurasi dengan *concatenating*, diperlukan sedikit modifikasi untuk melakukan *pre-training* graf pada teknik AMRBART. Token `<s>` yang digunakan pada AMRBART diganti menjadi kode bahasa yang digunakan. Bahasa Indonesia ditandai dengan token `<id>` dan Bahasa Inggris ditandai dengan token `<en>`. Kalimat berbahasa Indonesia dan Inggris dilinearisasi secara sejajar sebagai input *pre-training* maupun *fine-tuning*. Strategi *pre-training* dan *fine-tuning* untuk pelatihan graf model AMR *parsing* lintas bahasa tersebut dapat dilihat pada Table III.2. Output dari pelatihan tersebut merupakan graf AMR yang dilinearisasi, yaitu `<g>`$g_1, . . , g_m$`</g>`. Linearisasi graf menggunakan algoritma DFS (Bevilacqua dkk., 2021).



Gambar III.2: Pelatihan dengan skema *bilingual*.

Pembangunan data latih silver ketika *pre-training* dilakukan *masking* secara dinamis ketika melakukan training (Bai dkk., 2022). Fungsi *noising* dilakukan secara acak

Tabel III.1: Strategi *pre-training* (PT) dan *fine-tuning* (FT) untuk pelatihan graf model AMR *parsing* lintas bahasa pada skema pelatihan *bilingual*. t/g merupakan teks/graf *original*. $\hat{t}/\hat{g}$ merupakan teks/graf yang *noisy* (hasil *denoising*). $\bar{t}/\bar{g}$ merupakan teks/graf yang hilang. Semua output dari pelatihan ini adalah bentuk graf yang utuh.

Phase	Task	Input
PT	$\bar{t}\hat{g}2g$	<s> [mask] </s> <g> g_1 , .. [mask] .. , g_m </g>
	$t\hat{g}2g$	<s> a_1 , .. , a_n </s> <g> g_1 , .. [mask] .. , g_m </g>
	$\hat{t}\hat{g}2g$	<s> a_1 , .. [mask] .. , a_n </s> <g> g_1 , .. [mask] .. , g_m </g>
FT	$t\bar{g}2g$	<s> a_1 , .. , a_n </s> <g> g_1 , .. [mask] .. , g_m </g>

Tabel III.2: Strategi *pre-training* (PT) dan *fine-tuning* (FT) untuk pelatihan graf model AMR *parsing* lintas bahasa pada skema pelatihan *bilingual* dengan konfigurasi *concatenating*. t/g merupakan teks/graf *original*. $\hat{t}/\hat{g}$ merupakan teks/graf yang *noisy* (hasil *denoising*). $\bar{t}/\bar{g}$ merupakan teks/graf yang hilang. Semua output dari pelatihan ini adalah bentuk graf yang utuh.

Phase	Task	Input
PT	$\bar{t}_{ID}\bar{t}_{EN}\hat{g}2g$	<id> [mask] </id> <en> [mask] </en> <g> g_1 , .. [mask] .. , g_m </g>
	$t_{ID}\bar{t}_{EN}\hat{g}2g$	<id> a_1 , .. , a_{n_1} </id> <en> [mask] </en> <g> g_1 , .. [mask] .. , g_m </g>
	$\bar{t}_{ID}t_{EN}\hat{g}2g$	<id> [mask] </id> <en> b_1 , .. , b_{n_2} </en> <g> g_1 , .. [mask] .. , g_m </g>
	$t_{ID}t_{EN}\hat{g}2g$	<id> a_1 , .. , a_{n_1} </id> <en> b_1 , .. , b_{n_2} </en> <g> g_1 , .. [mask] .. , g_m </g>
	$t_{ID}\hat{t}_{EN}\hat{g}2g$	<id> a_1 , .. , a_{n_1} </id> <en> b_1 , .. [mask] .. , b_{n_2} </en> <g> g_1 , .. [mask] .. , g_m </g>
	$\hat{t}_{ID}t_{EN}\hat{g}2g$	<id> a_1 , .. [mask] .. , a_{n_1} </id> <en> b_1 , .. , b_{n_2} </en> <g> g_1 , .. [mask] .. , g_m </g>
	$\hat{t}_{ID}\hat{t}_{EN}\hat{g}2g$	<id> a_1 , .. [mask] .. , a_{n_1} </id> <en> b_1 , .. [mask] .. , b_{n_2} </en> <g> g_1 , .. [mask] .. , g_m </g>
FT	$t_{ID}\bar{t}_{EN}\bar{g}2g$	<id> a_1 , .. , a_{n_1} </id> <en> [mask] </en> <g> [mask] </g>
	$\bar{t}_{ID}t_{EN}\bar{g}2g$	<id> [mask] </id> <en> b_1 , .. , b_{n_2} </en> <g> [mask] </g>

untuk melakukan *masking* sebuah subgraf, 15% bagian untuk *masking* simpul (konsep), dan/atau 15% untuk *masking* sisi (relasi). Sebuah subgraf minimal memiliki satu simpul dan satu sisi. Setiap jenis *noising* tersebut memiliki kemungkinan yang bertambah setiap *step* pelatihan. Nilai kemungkinan tersebut dapat dilihat pada Persamaan (III.1). Nilai p merupakan kemungkinan dilakukan *noising*. Nilai t merupakan iterasi pelatihan dan T merupakan iterasi maksimal.

$$p = 0.1 + 0.75 \times t/T \tag{III.1}$$

Pelatihan AMR *parsing* akan dicoba pada beberapa *multilingual* model. Model-model yang akan dicoba adalah mBART, mT5, IndoBART, dan IndoT5.

Eksperimen AMR *parsing* juga akan dicoba dengan teknik AMR ensembling (Hoang dkk., 2021) dari hasil semua model-model tersebut.

Evaluasi dilakukan secara kuantitatif dan kualitatif. Metrik evaluasi yang digunakan untuk evaluasi kuantitatif adalah SMATCH. Evaluasi tersebut dilakukan dengan menghitung kemiripan graf AMR hasil prediksi dan graf AMR asli dari dataset test. Metode kualitatif yang digunakan adalah evaluasi divergensi translasi.

Pendekatan *translate-and-parse* (Uhrig dkk., 2021) digunakan sebagai baseline pada penelitian ini. Model mesin translasi yang digunakan adalah OPUS-MT dan model AMR *parsing* yang digunakan adalah AMRBART. Tahapan yang dilakukan adalah melakukan translasi dari kalimat berbahasa Indonesia menjadi Bahasa Inggris, lalu dilakukan AMR *parsing* menjadi graf AMR.

DAFTAR PUSTAKA

Bai, Xuefeng, Yulong Chen dan Yue Zhang (2022). "Graph Pre-training for AMR Parsing and Generation". Pada: *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, hal. 6001–6015.

Banarescu, Laura dkk. (2013). "Abstract meaning representation for sembanking". Pada: *Proceedings of the 7th linguistic annotation workshop and interoperability with discourse*, hal. 178–186.

Bevilacqua, Michele, Rexhina Blloshmi dan Roberto Navigli (2021). "One SPRING to rule them both: Symmetric AMR semantic parsing and generation without a complex pipeline". Pada: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 35. 14, hal. 12564–12573.

Blloshmi, Rexhina, Rocco Tripodi dan Roberto Navigli (2020). "XL-AMR: Enabling cross-lingual AMR parsing with transfer learning techniques". Pada: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, hal. 2487–2500.

Bojanowski, Piotr dkk. (2017). "Enriching word vectors with subword information". Pada: *Transactions of the association for computational linguistics* 5, hal. 135–146.

BPPT, PAN Localization - (2009). "Parallel Text Corpora, English Indonesian". Pada: URL: `http://digilib.bppt.go.id/sampul/p92-budiono.pdf`.

Cai, Deng dan Wai Lam (2020). "AMR Parsing via Graph-Sequence Iterative Inference". Pada: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, hal. 1290–1301.

Cai, Shu dan Kevin Knight (2013). "Smatch: an evaluation metric for semantic feature structures". Pada: *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, hal. 748–752.

Cettolo, Mauro dkk. (Des. 2017). "Overview of the IWSLT 2017 Evaluation Campaign". Pada: *Proceedings of the 14th International Conference on Spoken*

Language Translation. Tokyo, Japan: International Workshop on Spoken Language Translation, hal. 2–14. URL: https://aclanthology.org/2017.iwslt-1.1.

Conneau, Alexis dan Guillaume Lample (2019). "Cross-lingual language model pretraining". Pada: *Advances in neural information processing systems* 32.

Damonte, Marco dan Shay Cohen (2018). "Cross-lingual Abstract Meaning Representation Parsing". Pada: *16th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics (ACL), hal. 1146–1155.

Devlin, Jacob dkk. (2019). "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding". Pada: *Proceedings of NAACL-HLT*, hal. 4171–4186.

Dorr, Bonnie (1994). "Machine translation divergences: A formal description and proposed solution". Pada: *Computational linguistics* 20.4, hal. 597–633.

Feng, Fangxiaoyu dkk. (2022). "Language-agnostic BERT Sentence Embedding". Pada: *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, hal. 878–891.

Hoang, Thanh Lam dkk. (2021). "Ensembling Graph Predictions for AMR Parsing". Pada: *Advances in Neural Information Processing Systems* 34, hal. 8495–8505.

Ilmy, Adylan Roaffa dan Masayu Leylia Khodra (2020). "Parsing Indonesian Sentence into Abstract Meaning Representation using Machine Learning Approach". Pada: *2020 7th International Conference on Advance Informatics: Concepts, Theory and Applications (ICAICTA)*. IEEE, hal. 1–6.

Kingsbury, Paul R dan Martha Palmer (2002). "From treebank to propbank." Pada: *LREC*, hal. 1989–1993.

Lee, Young-Suk dkk. (Jul 2022). "Maximum Bayes Smatch Ensemble Distillation for AMR Parsing". Pada: *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Seattle, United States: Association for Computational

Linguistics, hal. 5379–5392. DOI: 10.18653/v1/2022.naacl-main.393. URL: https://aclanthology.org/2022.naacl-main.393.

Lewis, Mike dkk. (2020). "BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension". Pada: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, hal. 7871–7880.

Liu, Yinhan dkk. (2020). "Multilingual denoising pre-training for neural machine translation". Pada: *Transactions of the Association for Computational Linguistics* 8, hal. 726–742.

Mikolov, Tomás dkk. (2013). "Efficient Estimation of Word Representations in Vector Space". Pada: *1st International Conference on Learning Representations, ICLR 2013, Scottsdale, Arizona, USA, May 2-4, 2013, Workshop Track Proceedings*.

Pennington, Jeffrey, Richard Socher dan Christopher D Manning (2014). "Glove: Global vectors for word representation". Pada: *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, hal. 1532–1543.

Putra, Aditya Rachman (2022). "Pembangkitan Abstract Meaning Representation Lintas Bahasa dari Kalimat Berbahasa Indonesia". Tesis Program Magister. Institut Teknologi Bandung.

Radford, Alec dkk. (2019). "Language Models are Unsupervised Multitask Learners". Pada: *OpenAI blog, 1(8), 9*.

Raffel, Colin dkk. (2020). "Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer". Pada: *Journal of Machine Learning Research* 21, hal. 1–67.

Severina, Verena dan Masayu Leylia Khodra (2019). "Multidocument abstractive summarization using abstract meaning representation for Indonesian language". Pada: *2019 International Conference of Advanced Informatics: Concepts, Theory and Applications (ICAICTA)*. IEEE, hal. 1–6.

Tiedemann, Jörg dan Santhosh Thottingal (2020). "OPUS-MT Building open translation services for the World". Pada: *Proceedings of the 22nd Annual*

Conferenec of the European Association for Machine Translation (EAMT). Lisbon, Portugal.

Uhrig, Sarah dkk. (2021). "Translate, then Parse! A Strong Baseline for Cross-Lingual AMR Parsing". Pada: *Proceedings of the 17th International Conference on Parsing Technologies and the IWPT 2021 Shared Task on Parsing into Enhanced Universal Dependencies (IWPT 2021)*, hal. 58–64.

Vaswani, Ashish dkk. (2017). "Attention is all you need". Pada: *Advances in neural information processing systems* 30.

Wu, Ledell dkk. (2020). "Scalable Zero-shot Entity Linking with Dense Entity Retrieval". Pada: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, hal. 6397–6407.

Xue, Linting dkk. (2021). "mT5: A Massively Multilingual Pre-trained Text-to-Text Transformer". Pada: *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, hal. 483–498.

Xue, Nianwen dkk. (2014). "Not an interlingua, but close: Comparison of English AMRs to Chinese and Czech". Pada: *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, hal. 1765–1772.

Zhang, Sheng dkk. (2019). "AMR Parsing as Sequence-to-Graph Transduction". Pada: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, hal. 80–94.