

**PENGAUTOMASIAN PENGANALISISAN NOTA-NOTA KAKI
PENYATA KEWANGAN BERASASKAN PENGETAHUAN**

OLEH

ROHAYA AMAL BT WAN ABDUL RAZAK

**DISERTASI DIHANTAR SEBAGAI MEMENUHI SEBAHAGIAN
DARI KEPERLUAN PENGIJAZAHAN SARJANA
SAINS KOMPUTER
(DALAM BIDANG KECERDASAN BUATAN)**

JULAI 1991

PENGHARGAAN

Terlebih dahulu saya rasa bersyukur kepada Allah swt kerana telah memberikan saya kesabaran serta kekuatan untuk menghadapi segala cabaran, masalah serta rintangan yang telah saya hadapi semasa mengikuti kursus Sarjana ini.

Saya mengucapkan ribuan terima kasih terutama sekali kepada Yang Berbahagia Datuk Musa Mohamad selaku Naib Canselor Universiti Sains Malaysia yang telah membenarkan saya mengikuti kursus peringkat Sarjana Kecerdasan Buatan ini.

Saya juga mengucapkan ribuan terima kasih kepada penyelia yang saya hormati iaitu Profesor Madya Dr. Zaharin Yusoff yang telah mengambil berat serta telah memberikan pengajaran serta nasihat yang berguna kepada saya sepanjang saya mengikuti kursus ini. Nasihat-nasihat serta pengajaran beliau akan saya jadikan pedoman di dalam hidup saya.

Terima kasih juga saya ucapkan kepada Dr. Zahran Halim, Dr. Zaidah, Dr. Tong Loong Cheong, Dr. Nor Adnan Yahya, Cik Anna Bang dan Dr. Subramaniam dari Pusat Pengajian Pengurusan yang juga telah melibatkan diri mereka dengan kursus ini. Jasa-jasa serta nasihat-nasihat mereka tidak akan saya lupakan, terutama sekali kepada Dr. Zaidah yang telah banyak memberikan galakan serta dorongan kepada saya sepanjang tempoh kursus ini.

Akhir sekali, saya juga tidak lupa untuk mengucapkan terima kasih kepada keluarga yang saya sayangi, terutama sekali kepada suami saya yang juga telah memikul sebahagian daripada tanggungjawab saya, sepanjang saya mengikuti kursus Sarjana ini.

Sekian.

ROHAYA AMAL BT WAN ABDUL RAZAK

KANDUNGAN

BAB	MUKASURAT
BAB 1	(1 - 6)
1.0 PENGENALAN	1
1.1 Peranan Analisis Di Dalam Pemutusan Kewangan	1
1.2 Penyata-Penyata Kewangan Dan Nota-Nota Kaki	2
1.3 Objektif Penyelidikan	2
1.4 Pandangan Bagi Tugas Penyelidikan	4
 BAB 2	 (7 - 11)
2.0 ANALISIS MASALAH	7
2.1 Domain Masalah	7
2.2 Susutnilai ("Depreciation")	7
2.3 Nota-Nota Kaki Di Dalam Bentuk Bahasa Asli	8
2.3.1 Hubungan Semantik Asas: Proses, Objek dan Pemanfaat	8
2.3.2 Paras-Paras Tafsiran	9
2.4 Keperluan Pangkalan Pengetahuan	10
2.5 Pemetaan	10
 BAB 3	 (12-25)
3.0 PERWAKILAN PENGETAHUAN DAN PERWAKILAN TEKS INPUT	12
3.1 Perwakilan Yang Dipilih	12
3.2 Perwakilan Struktur Kerangka	12
3.3 Tatatanda Dan Simbol Yang Digunakan	14
3.4 Rekabentuk Struktur Kerangka	15
3.4.1 Kerangka Aset	16
3.4.2 Kerangka Depn_amortn	20
3.4.2.1 Slot Aset_type Dan Slot Process	20
3.4.2.2 Slot Method_used	20
3.4.2.3 Slot Min_annual_rate, Max_annual_rate, Slot Min_depn_value, Max_Depn_value, Slot Owned_by	21
3.4.3 Kerangka Business_orgn	22
3.5 Kesimpulan Bagi Rekabentuk	23
3.6 Perwakilan Struktur-M	23
3.7 Struktur-M dan Peranan Semantik	24
 BAB 4	 (26-47)
4.0 REKABENTUK PEMETAAN	26
4.1 Pemetaan	26
4.2 Pengecaman Kata Kunci	26
4.3 PRED atau Proses Utama	27
4.4 Pengecaman Nota Kaki Domain Susutnilai	28

4.5	Strategi Pemetaan	2 8
4.6	Kawalan Pemetaan	3 9
4.6.1	Pengaktifan Kerangka	3 9
4.6.2	Pemilihan Kerangka Anak Di Dalam Pautan IS_A	4 2
4.6.3	Penginstantasian Slot	4 2
4.6.4	Kesimpulan Strategi Kawalan Pemetaan	4 3
4.6.5	Kawalan Proses Pemetaan Bagi Domain Masalah	4 4
BAB 5		(48-52)
5.0	IMPLEMENTASI	48
5.1	Petua-Petua Pemetaan	48
5.2	Alat-Alat Implementasi Dan Format Input	48
5.3	Output Penyelidikan	4 9
5.4	Proses Pengaliran	5 1
BAB 6		(53-54)
6.0	PERBINCANGAN DAN TUGAS MASA HADAPAN	5 3
6.1	Perbincangan	5 3
6.2	Masalah-Masalah	5 3
6.3	Tugas-Tugas Masa Hadapan	5 3
6.4	Sumbangan Penyelidikan	5 4
BAB 7		
7.0	KESIMPULAN	5 5

Apendiks A	Ayat-Ayat DI Dalam Korpus
Apendiks B	Perwakilan Struktur-m Ayat-Ayat DI Dalam Korpus
	Perwakilan Logik Ayat-Ayat DI Dalam Korpus
Apendiks C	Petua-Petua Pemetaan
Apendiks D	Output Penyelidikan

ABSTRACT

THE AUTOMATION OF KNOWLEDGE BASED ANALYSIS OF FINANCIAL STATEMENT FOOTNOTES

The main objective is to arrive at a design of a pedagogical model for the transformation of financial statement footnotes (in textual form) into frame-based knowledge representation. The implementation representation adopted is the Framelog representation [Zahran 87, 89, 90], which is basically a combination of two knowledge formalisms, i.e. frame structures and the logic formalism. The unstructured information in the financial statement footnotes is captured and extracted with natural language processing techniques, in particular using the analyser in the JEMAH system [Tong 88], which results in a linguistic abstract structure called the m-structure [Zaharin 87]. The work done for this dissertation completes the transformation by interpreting the m-structure to obtain the appropriate Framelog representation. The knowledge representation of the financial statement footnotes finally produced is planned to be used in financial analysis and interpretation systems, for example CoPE [Zaidah 89].

ABSTRAK

Matlamat utama di sini adalah untuk memperolehi suatu rekabentuk model yang bersifat pedagogi bagi transformasi/ penjelmaan nota kaki penyata kewangan (dalam bentuk teks) ke dalam bentuk perwakilan pengetahuan berasaskan kerangka. Perwakilan implementasi yang dipilih untuk tugas ini ialah perwakilan Framelog [Zahran 87, 89, 90], yang pada dasarnya merupakan gabungan dua formalisma pengetahuan, iaitu struktur kerangka dan formalisma logik. Maklumat dalam nota-nota kaki penyata kewangan yang tidak berstruktur itu diperolehi dan diekstrak melalui kaedah-kaedah pemprosesan bahasa asli, khususnya dengan menggunakan penganalisis di dalam sistem JEMAH [Tong 88], yang menghasilkan struktur abstrak linguistik yang dipanggil struktur-m [Zaharin 87]. Hasil tugas disertasi ini melengkapkan transformasi yang diperlukan dengan menafsirkan struktur-m untuk memperolehi perwakilan Framelog yang berkaitan. Perwakilan pengetahuan bagi nota-nota penyata kewangan yang terhasil, dijangka akan digunakan di dalam sistem-sistem analisis dan interpretasi, umpamanya sistem CoPE [Zaidah 89].

BAB 1

1.0 PENGENALAN

1.1 Peranan Analisis Kewangan Di Dalam Pemutusan Kewangan ('Financial Decision Making')

Analisis kewangan ialah sistem pemprosesan maklumat yang direkabentuk untuk menyediakan dan membekalkan data khusus untuk orang yang terlibat dalam usaha membuat keputusan [Lev 74]. Berbanding dengan sistem analisis tradisional, sistem analisis cara moden tidak lagi terhad kepada data perakaunan yang terdapat di dalam penyata-penyata kewangan. Kini maklumat-maklumat diperolehi daripada penyata-penyata kewangan (bersama dengan nota-nota kakinya ('footnotes')), dan juga daripada data yang belum dilaporkan, seperti nilai pasaran sesuatu harta dan ramalan penghasilan masa depan. Usaha analisis juga melibatkan data bukan perakaunan, seperti harga-harga jaminan keselamatan. Seseorang yang terlibat di dalam pemutusan perlu menumpukan perhatian kepada keadaan ekonomi sesuatu syarikat dan juga peramalan masa depannya.

Menurut [Mui 87], pemutusan kewangan terdiri daripada dua fasa, iaitu fasa pengenalan ('familiarisation phase') dan fasa penaakulan ('reasoning phase'). Fasa-fasa ini dapat diperlihatkan seperti langkah-langkah berikut:

- (1) pengimbasan nilai-nilai persekitaran dan latarbelakang untuk mengenalpasti item-item kunci ('key items'), seperti "sales" dan "net income".
- (2) pencapaian templet kewangan ('financial templates') bagi syarikat-syarikat atau perindustrian tertentu daripada ingatan jangka panjang (sebagai contoh, "high tech company" atau "late recessionary industry").
- (3) tindakan instantasi ("instantiations") sesuatu templet dengan sesuatu maklumat.
- (4) penilaian, atau pemutusan secara keseluruhan.

IFasa pengenalan boleh dipadankan dengan langkah pertama di atas, yang terdiri daripada tugas-tugas analisis. Perhatian akan ditumpukan kepada fasa tersebut di dalam disertasi ini. Fasa penaakulan berpadanan dengan langkah-langkah lain di atas, yang tidak akan diperbincangkan lagi seterusnya.

Menurut [Fraser 85], objektif tugas analisis mungkin berubah bergantung kepada perspektif pengguna penyata kewangan dan jenis soalan yang mungkin dikemukakan. Sebagai contoh, jenis soalan yang dikemukakan oleh pengkredit dalam analisis kredit ternyata berbeza daripada soalan-soalan yang dikemukakan oleh pelabur-pelabur dalam analisis pelaburan. Pengkredit lebih mengambil berat tentang kemampuan peminjam membayar balik faedah dan prinsipal bagi wang yang dipinjam, manakala pelabur pula berusaha membuat jangkaan pengaliran pendapatan masa depan syarikat supaya sesuatu nilai dapat diikaitkan dengan keselamatan dalam pertimbangan terhadap pembelian atau "liquidation".

Disertasi ini membincangkan tugas penganalisan maklumat yang pada dasarnya terdapat di dalam bentuk bahan-bahan bercetak, khususnya penyata-penyata kewangan bersama dengan nota-nota kakinya.

1.2 Penyata-Penyata Kewangan Dan Nota-Nota Kaki

Pada amnya, kebanyakan penyata kewangan yang terdapat di dalam laporan tahunan bagi sesuatu syarikat mengandungi empat jenis maklumat kewangan. Mengikut [Fraser 85], maklumat tersebut boleh dikelaskan sebagai:-

- (1) kunci kira-kira ("balance sheet") yang menunjukkan kedudukan kewangan syarikat itu;
- (2) penyata pendapatan ("statement of income or earnings");
- (3) penyata penyimpanan pendapatan ("statement of retained earnings");
- (4) penyata segala perubahan dalam kedudukan kewangan yang dapat memberi maklumat tentang aktiviti-aktiviti kewangan dan pelaburan dalam jangkamasa tertentu bagi syarikat itu.

Nota-nota kaki yang disertakan bersama dengan penyata kewangan, adalah amat penting dan perlu dibaca supaya penyata itu dapat difahami sepenuhnya. Nota-nota ini mungkin mempunyai maklumat tambahan yang tidak wujud di dalam penyata asal, seperti yang berkaitan dengan "depreciation", "pensions", "contingent liabilities" , dsb.

Maklumat yang diperolehi daripada penyata kewangan bersama dengan nota kaki seperti yang disebutkan di atas digunakan di dalam analisis untuk menyediakan data yang diperlukan untuk membuat ramalan dan juga keputusan perniagaan terhadap sesuatu syarikat.

1.3 Objektif Penyelidikan

Jelas bahawa sesuatu analisis memerlukan maklumat daripada beberapa mukasurat penyata dan nota kaki agar maklumat yang mencukupi dapat diperolehi. Tambahan pula, maklumat ini perlu difahami dan dibentangkan dalam bentuk yang sesuai ("convivial"). Oleh demikian, masa dan usaha dalam tugas analisis mungkin dikurangkan dengan kadar yang banyak sekiranya tugas analisis penyata kewangan ini dapat diautomasikan ("automated financial statement analysis").

Perkara-perkara yang akan diperihalkan di dalam disertasi ini ialah yang berkenaan dengan kaedah implementasi bagi pengautomasian analisis penyata kewangan, terutamanya melalui penggunaan teknik-teknik kecerdasan buatan. Kini terdapat beberapa sistem analisis penyata kewangan yang sedia ada, umpamanya sistem FSA [Mui 87]. Sistem ini merupakan sistem kecerdasan buatan yang dibina oleh Arthur Anderson & Company di bawah kontrak dengan SEC sebagai fasa permulaan di dalam projek EDGAR (Electronic Data Gathering, Analysis and Retrieval). Sistem ini berorientasikan objek dan strukturnya dimodelkan mengikut struktur pengetahuan domain perakaunan. FSA telah cuba berusaha membina penganalisisnya sendiri untuk pengendalian nota kaki, tetapi yang dihasilkan tidak begitu menyeluruh ("

partially built") oleh kerana ia didapati sesuai hanya untuk domain yang dikendalikan oleh FSA sahaja dan bukan untuk domain-domain penerapan lain. Penyelidikan kami menggunakan suatu penganalisis yang sedia ada, diperolehi daripada sistem JEMAH [Tong 88]. JEMAH merupakan sebuah sistem penterjemahan berkomputer yang mempunyai modul penganalisis yang berasingan. Jika dibandingkan dengan penganalisis yang dibina oleh FSA, penganalisis JEMAH boleh dikatakan lebih umum oleh kerana ia

- sedia digunakan;
- tidak bersandaran kepada sesuatu domain tertentu;
- tidak bersandaran kepada sesuatu bahasa tertentu.

Tambah pula, modul analisis bagi bahasa Inggeris sudah sedia terbina di dalam sistem penterjemahan bahasa Inggeris kepada bahasa Malaysia. Modul ini boleh terus digunakan untuk kajian disertasi ini. Kebaikan-kebaikan ini dapat mengatasi masalah dalam memperolehi penganalisis yang mungkin berkaitan rapat dengan domain masalah dan juga bahasa yang digunakan, seperti yang terdapat dalam FSA.

Kami bantangkan di sini suatu usaha ke arah pembinaan suatu sistem yang akan mengautomasikan tugas penganalisan nota kaki penyata kewangan bagi membekalkan serta menyediakan data yang berguna di dalam usaha membuat keputusan ("decision making"). Hasil output sistem ini ditujukan sebagai bahan asas untuk kegunaan sistem CoPE [Zaidah 89]. CoPE merupakan sistem pakar yang bertindak secara interaktif di dalam usaha menyediakan panduan bagi penilaian prestasi sesuatu syarikat.

Tugas analisis yang kami maksudkan di dalam disertasi ini merupakan usaha penukaran atau transformasi data penyata kewangan dan nota kakinya (yang tidak berstruktur), ke dalam bentuk maklumat yang lebih sesuai digunakan. Walau bagaimana pun, sebagaimana yang telah dinyatakan di dalam bahagian 1.1, objektif muktamad bagi analisis ini mungkin berubah mengikut keperluan pengguna dan juga soalan-soalan yang dikemukakan (khususnya yang dikemukakan oleh CoPE, bagi kes kami).

Tugas penyelidikan kami merupakan sebahagian daripada tugas sistem CoPE. Sistem CoPE ("Company Performance Evaluation") ialah sebuah sistem pangkalan pengetahuan yang lengkap atau komprehensif, dan direkabentuk khusus untuk menilai kemungkinan pengintegrasian beberapa komponen yang saling berinteraksi, umpamanya sistem-sistem pangkalan pengetahuan, sistem-sistem bantuan pengambilan keputusan dan sistem-sistem pemprosesan data. Pada amnya, CoPE merupakan sebuah sistem bantuan keputusan yang cerdas dan interaktif, yang membantu pembuat keputusan (manusia) menghasilkan keputusan yang komprehensif dan konsisten, dalam usaha menilai prestasi sesuatu syarikat. Penyelidikan ke arah menyediakan "natural language front end" [Yeoh 89] serta suatu pengelasan kemasukan penyata kewangan dengan menggunakan kerangka ("frames") [Lim 89] (setiap satu merupakan sebahagian daripada CoPE), telah pun dilaksanakan dengan hasil yang agak memuaskan.

Bahagian-bahagian berikut memerihalkan pandangan umum ("overview") bagi tugas penyelidikan yang diperbincangkan.

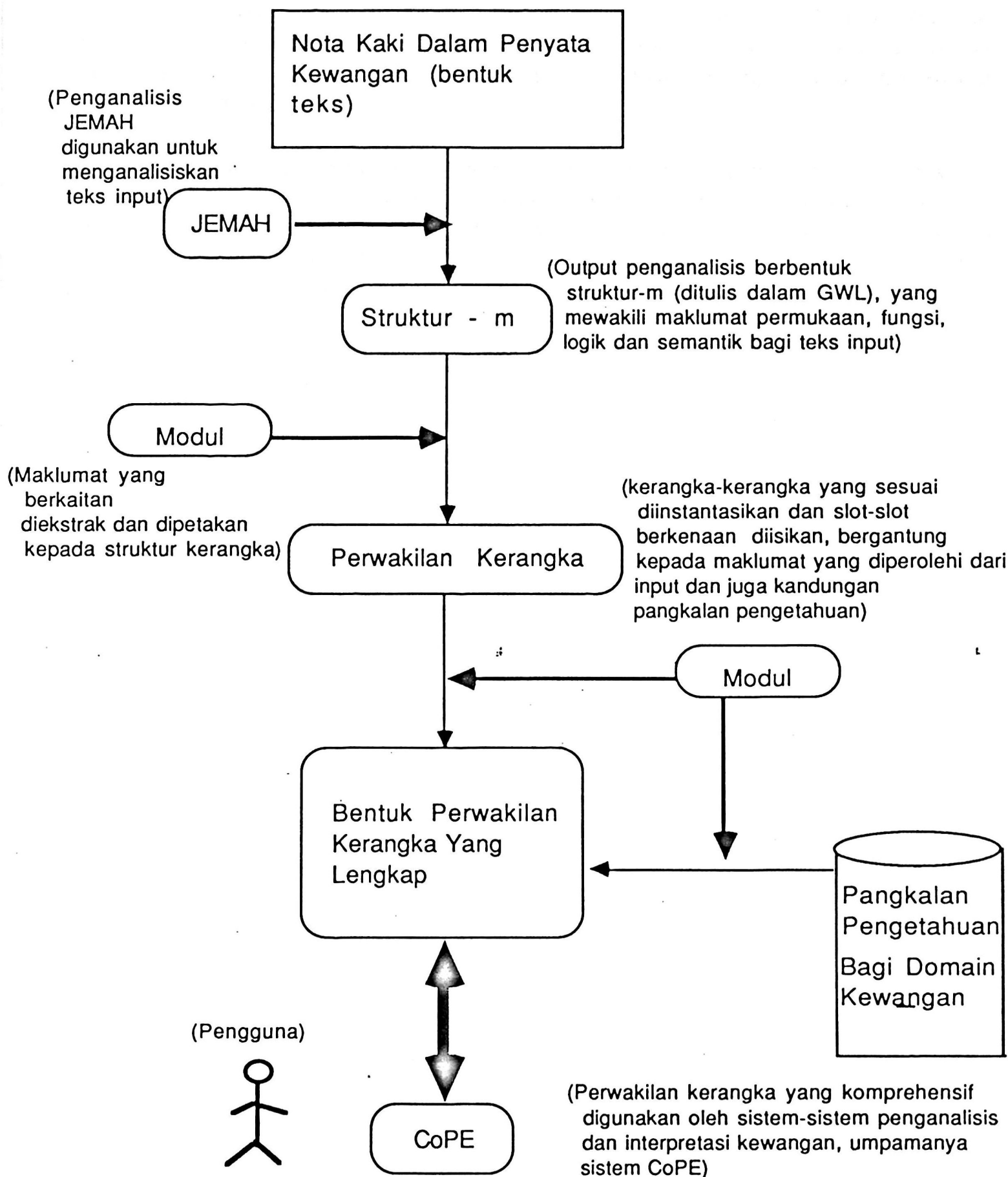
1.4 Pandangan Umum Terhadap Tugas Penyelidikan

Pandangan secara keseluruhan bagi tugas yang dilaksanakan di dalam disertasi ini dirumuskan di dalam gambarajah 1.1. Sebagaimana yang dinyatakan, input merupakan set nota kaki yang disertakan bersama dengan penyata kewangan di dalam laporan tahunan sesuatu syarikat. Sistem yang dicadangkan menganalisis nota-nota kaki tersebut untuk menyediakan output yang merupakan perwakilan kerangka yang komprehensif sebagai perwakilan maklumat yang terkandung di dalam input yang seterusnya akan digunakan di dalam analisis kewangan di dalam sistem CoPE.

Dengan lebih terperinci, nota-nota kaki berbentuk teks dianalisis dengan menggunakan penganalisis daripada sistem JEMAH [Tong 88] (yang merupakan sebuah sistem penterjemahan berkomputer). Hasil output penganalisis ini berbentuk struktur pohon berhias ("annotated tree") mengikut perwakilan konsep struktur-m ("m-structure") [Vauquois 78, Zaharin 87]. Struktur tersebut merupakan perwakilan bagi maklumat yang merangkumi permukaan ("surface"), fungsi ("functional"), logik ("logical") dan semantik ("semantic") bagi teks input. Di dalam JEMAH, struktur-m diwakili di dalam formalisma GWL ("Grammar Writing Language"), yang dijadikan sebagai input bagi tugas kami. Perwakilan abstrak GWL berbentuk senarai pembolehubah yang tersirat, seakan-akan struktur LISP.

Di dalam modul utama, maklumat-maklumat yang berkaitan dengan penganalisan kewangan akan diekstrakkan ("extract") daripada input struktur-m, khususnya yang didapati di dalam komponen-komponen logik dan semantik. Komponen permukaan ("surface") dan fungsi ("functional") lebih merupakan maklumat sintaksis yang biasanya hanya diperlukan untuk melaksanakan analisis linguistik. Maklumat yang diperolehi seterusnya diubah bentuk ke dalam struktur yang berasaskan kerangka [Fikes 85]. Transformasi ini melibatkan tindakan instantiasi ("instantiation") kerangka-kerangka yang sesuai serta pengisian slot-slot yang berpadanan bagi struktur-struktur kerangka umum domain kewangan yang disediakan. Selain merekabentuk dan mengimplementasikan modul-modul untuk pelaksanaan transformasi utama ini, sumbangan disertasi ini juga merangkumi rekabentuk struktur kerangka umum yang disebutkan di atas. Bagaimanapun, kami telah menghadkan domainnya kepada domain susutnilai ("depreciation") sahaja, oleh kerana tujuan utama projek ini adalah untuk menguji kesauran kaedah yang dicadangkan sahaja, dan bukan untuk membina sebuah sistem yang menyeluruh.

Sesudah langkah transformasi di atas, struktur kerangka umum tersebut mungkin belum lagi diinstantiasikan sepenuhnya, tetapi sekurang-kurangnya ia mengandungi kesemua maklumat tepat yang boleh didapati di dalam input struktur-m (yang mewakili maklumat daripada teks input nota kaki kewangan yang asal). Untuk melengkapkan struktur kerangka ini, pangkalan pengetahuan bagi domain kewangan dirujuk, dan proses pencapaian ini juga memerlukan rekabentuk dan implementasi modul-modul yang sesuai.



GAMBARAJAH 1.1. PANDANGAN UMUM TUGAS PENYELIDIKAN

Semua struktur kerangka yang dinyatakan di sini direkabentuk mengikut subdomain bagi domain pengetahuan kewangan yang disebutkan terlebih awal, iaitu susutnilai ("depreciation"). Dari segi implementasi, Framelog [Zahran 87,90] digunakan sebagai alat kejuruteraan pengetahuan. Framelog merupakan sebuah sistem pakar umum yang menawarkan dua jenis formalisma, iaitu kerangka dan logik (khususnya bahasa pemrograman PROLOG), dan tiga jenis "facet", iaitu "question facet", "if-needed facet" dan "default facet".

Pangkalan pengetahuan bagi domain kewangan yang digunakan di sini ialah yang terdapat di dalam sistem CoPE. Pengetahuan ini juga diwakilkan dengan menggunakan formalisma kerangka. Struktur kerangka terakhir (yang dihasilkan daripada peringkat transformasi dan seterusnya melalui rujukan kepada pangkalan pengetahuan) kini sedia digunakan untuk peringkat tafsiran dalam analisis kewangan. Bagi kami, tugas tersebut dikendalikan oleh sistem CoPE.

BAB 2

2.0 ANALISIS MASALAH

2.1 Domain Masalah

Setiap tugas analisis nota kaki penyata kewangan perlu mengendalikan banyak jenis subdomain kewangan yang mungkin terdapat di dalam nota kaki yang sama. Subdomain-subdomain ini boleh terdiri daripada polisi perakaunan, asas penyatuan ("basis of consolidation"), penilaian pelaburan ("valuation of investments"), stok dan sebagainya. Di dalam bidang tugas penyelidikan ini, hanya domain susutnilai ("depreciation") (iaitu suatu bahagian daripada polisi perakaunan) akan kami titikberatkan. Pilihan ini dibuat disebabkan masa yang agak terhad dan juga berdasarkan hasrat kami untuk menganalisis sesuatu masalah itu dengan mendalam supaya dapat menghasilkan satu contoh pedagogi yang baik dan dengan itu dapat diperluaskan seterusnya kepada domain-domain penerapan yang lain.

Oleh kerana kebanyakan syarikat di Malaysia menyediakan laporan tahunan mereka di dalam bahasa Inggeris, kami menumpukan perhatian hanya kepada nota-nota kaki di dalam bahasa tersebut. Namun begitu, kami percaya cadangan-cadangan serta pendapat-pendapat yang dikemukakan di dalam disertasi ini mungkin dilaksanakan juga kepada nota-nota kaki di dalam Bahasa Malaysia ataupun bahasa-bahasa lain.

Kami mulakan dengan penjelasan ringkas beberapa perkara dan pertimbangan asas yang akan digunakan di dalam penyelidikan kami, dan seterusnya akan dianggap sudah diketahui.

2.2 Susutnilai ("Depreciation")

Berkaitan dengan domain tumpuan kami, [Hayward 86] menyatakan bahawa domain susutnilai di dalam nota-nota kaki pada amnya mengandungi maklumat seperti berikut:-

- (1) kelas-kelas harta yang disusutnilaikan ("depreciated assets");
- (2) kelas -kelas harta yang tidak disusutnilaikan ("non depreciated assets");
- (3) kaedah-kaedah susutnilai ("depreciation methods") yang digunakan di dalam proses susutnilai untuk setiap kelas harta;
- (4) kadar-kadar prinsipal tahunan ("principal annual rates") yang digunakan untuk menghitung nilai yang disusutkan bagi setiap kelas harta yang disusutnilaikan.

Tumpuan akan diberikan kepada maklumat-maklumat (1) hingga (4) yang akan diperihalkan dengan lebih mendalam di dalam bab 3.

2.3 Nota-Nota Kaki Di Dalam Bentuk Bahasa Asli

Penyelidikan kami berdasarkan sebuah korpus kecil yang dipetik daripada beberapa nota-nota kaki penyata kewangan bagi beberapa organisasi berlainan. Di dalam korpus itu, lima belas ayat telah dipetik untuk dikaji. Ayat-ayat di dalam korpus tersebut disenaraikan di dalam Apendiks A. Namun demikian, buat sementara ini, hanya dua belas ayat sahaja (iaitu ayat(1) hingga (12)) yang merupakan ayat mudah telah kami gunakan di dalam penyelidikan. Walaupun ayat-ayat tersebut sedikit dan mudah, kami percaya kaedah yang digunakan membentuk suatu kaedah pedagogi yang dapat diperkembangkan kepada ayat yang lebih kompleks dan juga kepada korpus yang lebih besar dan lebih umum.

2.3.1 Hubungan Semantik Asas : Proses, Objek dan Pemanfaat

Bagi sesuatu ayat, kajian linguistik klasik menyatakan bahawa terdapat sekurang-kurangnya tiga peranan penting yang mungkin dimainkan oleh sesuatu rangkai kata ("phrase") di dalam struktur ayat atau klausa ("clause"), khususnya proses ("process"), objek ("object") dan pemanfaat ("beneficiary"). Peranan-peranan ini, yang dikenali sebagai hubungan semantik, akan dijadikan asas bagi pendekatan analisis ayat yang digunakan.

Proses merupakan suatu perbuatan ("action") yang dapat mengubah keadaan sesuatu objek (yang dilakukan perbuatan ke atasnya), manakala pemanfaat adalah merupakan pihak ketiga yang menerima manfaat daripada perbuatan itu. Sebagai contoh, di dalam ayat

'John gives a book to Mary',

give merupakan suatu proses, **a book** merupakan objek dan **Mary** merupakan pemanfaat.

Walaupun tidak semestinya benar bagi semua kes, sesuatu proses itu biasanya dihuraikan sebagai suatu perbuatan manakala objek dan pemanfaat dihuraikan sebagai nama, benda, manusia, dan sebagainya.

Berikut diberikan beberapa contoh tambahan. Di dalam ayat-ayat ini, proses ditunjukkan dalam **bold**, object ditunjukkan dalam *italics* dan pemanfaat ditunjukkan dengan garisan.

(i) *The principal annual rate* **used** for long leasehold land property is 2% to 4%.

(ii) *Plant and machinery under installation* **are not depreciated**.

(iii) *No depreciation* **is charged** for equipment and projects not commissioned.

(iv) *Depreciation* **is not provided** on freehold land.

(v) On other assets, *depreciation* **is computed** principally on the straight line method based on the estimated useful lives of the related assets.

(vi) *All other assets are depreciated on a straight line basis so as to write off their costs or valuations over their estimated useful lives.*

Sesuatu proses biasanya mempunyai suatu objek dan mungkin juga suatu pemanfaat. Sekiranya ayat itu di dalam ragam aktif, objek dijumpai selepas proses, tetapi biasanya di sebaliknya bagi ayat ragam pasif. Pemanfaat pula biasanya dijumpai selepas proses didahului suatu sendi nama yang sesuai. Di dalam ayat pasif, pemanfaat mungkin diletakkan sebelum proses.

Objek, proses dan pemanfaat memainkan peranan asas dalam membolehkan kita mengekstrakkan maklumat daripada sesuatu ayat. Perkara ini akan dibincangkan selanjutnya di dalam bahagian-bahagian seterusnya. Buat masa ini, eloklah disebutkan bahawa kejayaan analisis sesuatu ayat bergantung hampir sepenuhnya kepada pengecaman ("recognition") peranan-peranan asas tersebut.

2.3.2 Paras-Paras Tafsiran

Nota-nota kaki ialah ayat-ayat dalam bahasa asli dan dengan itu dapat dipandang dari segi sekurang-kurangnya empat paras tafsiran, khususnya permukaan ("surface"), fungsi ("functional"), logik dan semantik. Tafsiran permukaan memberi pandangan sesuatu ayat dari segi dekomposisinya ("decomposition") ke dalam kelas-kelas morfosintaktik ("morphosyntactic classes") [seperti **noun phrase** (NP), **verbal kernel** (VK) dan **prepositional phrase** (PP), dsb]. Tafsiran fungsi pula menunjukkan fungsi-fungsi sintaksis bagi rangkai-rangkai kata ("phrases") di dalam ayat itu [seperti **governor** (GOV), **objek** (OBJ) dan **subjek** (SUBJ), **circumstantial** (CIRC)]. Tafsiran semantik memberikan makna semantik kepada sesuatu ayat itu [seperti **agent**, **patient**, **instrument** (INST), **method** dsb]. Tafsiran logik pula memberikan makna logik kepada sesuatu ayat [seperti **Arg0**, **Arg1**, **Arg2**, **Gra0**, **Gra1**, **Trl01**, **meta-predicate** (META-PRED) dsb]. Sebagai contoh, paras-paras tafsiran di dalam ayat "The boy hits the dog with a stick" ialah,

permukaan : "The boy hits the dog with a stick"
NP VK NP PP

fungsi : "The boy hits the dog with a stick"
SUBJ GOV OBJ CIRC

logik : "The boy hits the dog with a stick"
ARG0 PRED ARG1 META-PRED

semantik : "The boy hits the dog with a stick"
AGENT PRED PATIENT INST

PRED di sini merupakan suatu proses seperti yang telah pun diperihalkan di dalam bahagian 2.3.1, PATIENT pula merupakan suatu objek, dan mengikut tafsiran semantik BENEF pula merupakan pemanfaat. Sila rujuk kepada [Zaharin 89] untuk mengetahui dengan lebih lanjut mengenai paras-paras tafsiran itu.

Pada dasarnya, tafsiran permukaan dan fungsi bagi sesuatu ayat menakrifkan kegramatisan ayat itu ("grammaticality of a sentence") manakala tafsiran logik dan semantik memberi makna asas ayat tersebut.

Tugas analisis harus mendapatkan tafsiran untuk paras logik dan semantik, kerana kami perlu menitikberatkan pengecaman peranan-peranan seperti **PRED**, **PATIENT** dan **BENEF**. Akan tetapi, analisis secara terus daripada teks input kepada paras-paras logik dan semantik tersebut amat sukar dilaksanakan dan tugasnya tidak dapat ditetapkan dengan persis. Oleh itu, kemungkinan paras-paras permukaan dan fungsi (yang lebih mudah dilengkapi) perlu dirujuk supaya dapat membantu di dalam pencapaian paras-paras logik dan semantik [Vauquois 78].

Perkara ini akan diperbincangkan lagi apabila kami memerihalkan proses analisis nanti. Bahagian berikutnya membincangkan keperluan pangkalan pengetahuan dalam penyelidikan ini.

2.4 Keperluan Pangkalan Pengetahuan

Menurut [Bratko 86], pangkalan pengetahuan adalah penting kerana ia mengandungi pengetahuan tentang sesuatu domain penerapan (bagi kes kami, domain kewangan). Pengetahuan itu diwakili oleh fakta-fakta mudah yang memerihalkan sesuatu domain, petua-petua ("rules") yang memerihalkan perhubungan-perhubungan atau fenomena yang didapati di dalam domain tersebut, dan kemungkinan juga heuristik dan cadangan-cadangan bagi menyelesaikan masalah-masalah yang berkaitan dengan domain itu.

Penyelidikan kami hanya menitikberatkan pangkalan pengetahuan yang mengandungi pengetahuan mengenai domain susutnilai, yang merupakan suatu subdomain kewangan. Perwakilan pengetahuan kami adalah di dalam bentuk struktur kerangka [Fikes 85]. Ia dipilih kerana CoPE menggunakan perwakilan pengetahuan yang sama. Rekabentuk struktur kerangka tersebut dilaksanakan oleh kami sendiri dan akan diperbincangkan dengan terperinci di dalam bab 3. Struktur kerangka ini kosong pada mulanya, iaitu belum diinstantasikan dan oleh itu belum sedia digunakan di dalam sesuatu penerapan.

Pemetaan di antara teks input dan struktur kerangka di atas perlu ditakrifkan, dan perkara ini akan diperbincangkan secara ringkas pada bahagian berikut. Hasil daripada pemetaan ini akan menyebabkan struktur kerangka tersebut diinstantasikan (mungkin tidak sepenuhnya) berdasarkan maklumat-maklumat yang diperolehi daripada teks input. Struktur kerangka yang telah diinstantasikan akan membentuk sebahagian daripada pangkalan pengetahuan CoPE dan seterusnya dapat digunakan sewajarnya.

2.5 Pemetaan

Pemetaan ("mapping") yang dimaksudkan dilaksanakan di antara teks input nota-nota kaki dengan struktur kerangka yang telah direkabentuk. Pemetaan ini bermula dengan analisis teks input untuk memperolehi struktur-m yang berkaitan yang mengandungi empat paras tafsiran seperti yang diperbincangkan di dalam bahagian 2.3.2. Tafsiran tersebut seterusnya digunakan di dalam proses

pemetaan (khususnya berdasarkan maklumat-maklumat logik dan semantik) seperti yang dirumuskan berikut :-

(1) Maklumat-maklumat yang berguna dan berkaitan dengan sesuatu domain harus diekstrakkan;

(2) Struktur kerangka yang direkabentuk diinstantasikan bergantung kepada maklumat-maklumat yang diekstrak.

Di dalam proses (1), maklumat-maklumat yang dikehendaki bergantung kepada maklumat yang diperolehi daripada domain susutnilai (seperti yang telah diperbincangkan di dalam bahagian 2.2).

Di dalam proses (2), kawalan serta strategi pemetaan perlu dikenalpasti untuk melaksanakannya.

Rekabentuk pemetaan ini merupakan proses utama penyelidikan kami, dan akan diperbincangkan dengan lebih terperinci di dalam bab 4. Sebelum itu, berikut disenaraikan perkara-perkara yang perlu ditelitikan untuk menjayakan usaha ini:-

(1) Rekabentuk struktur kerangka berasaskan pengetahuan kewangan, tetapi hanya untuk domain susutnilai.

(2) Pengkajian mengenai empat jenis tafsiran, iaitu permukaan, fungsi, logik dan semantik, yang dihasilkan oleh proses analisis teks input.

(3) Rekabentuk pemetaan (di antara teks input dengan struktur kerangka yang direkabentuk), yang merupakan kajian utama penyelidikan.

BAB 3

3.0 PERWAKILAN PENGETAHUAN DAN PERWAKILAN TEKS INPUT

3.1 Perwakilan Yang Dipilih

Dalam pembinaan suatu pangkalan pengetahuan, pemilihan perwakilan merupakan perkara yang amat penting oleh kerana perwakilan inilah yang perlu menyokong penghuraian maklumat dan juga kaedah pencapaian. Terdapat beberapa jenis perwakilan pengetahuan yang digunakan pada amnya, seperti petua, kerangka, logik, rangkaian semantik, dsb.

Kajian kami melibatkan dua jenis perwakilan, iaitu struktur kerangka dan struktur-m. Sebagaimana yang telah dinyatakan dalam bab yang lepas, struktur kerangka digunakan untuk mewakili pengetahuan, manakala struktur-m digunakan untuk mewakili aspek-aspek linguistik bagi maklumat yang terkandung di dalam nota kaki. Struktur-m ini di sini boleh dipandang sebagai struktur interim bagi analisis nota kaki dalam bentuk teks kepada perwakilannya dalam bentuk kerangka.

Kami mulakan perbincangan dengan perwakilan struktur kerangka pada amnya dan seterusnya rekabentuk struktur kerangka untuk domain susutnilai.

3.2 Perwakilan Struktur Kerangka

Sebagaimana yang telah dinyatakan dalam bab-bab yang lepas, kami menggunakan struktur kerangka untuk mewakili pengetahuan kewangan bagi domain susutnilai. Struktur kerangka dipilih khususnya kerana CoPE menggunakan perwakilan yang sama di dalam pangkalan pengetahuannya. Oleh kerana tugas penyelidikan kami merupakan sebahagian daripada rangka tugas pembangunan CoPE, maka pemilihan struktur ini akan dapat menjimatkan masa dan tenaga yang diperlukan. Tambah pula, memang terdapat beberapa kebaikan dengan penggunaan struktur kerangka sebagai perwakilan pengetahuan.

Secara amnya, sesuatu kerangka [Fikes 85] memerihalkan suatu kelas objek. Ia mengandungi satu koleksi slot yang memerihalkan aspek-aspek atau sifat-sifat ("properties") objek tersebut. Subdomain-subdomain kewangan (termasuk domain susutnilai) boleh diuraikan sebagai objek-objek atau kelas-kelas yang mempunyai beberapa sifat. Sebagai contoh, domain susutnilai boleh diuraikan sebagai suatu kelas yang mempunyai sifat-sifat seperti kadar-kadar minimum dan maksimum, kaedah-kaedah susutnilai ("depreciation methods") yang digunakan, dan sebagainya. Sifat-sifat ini akan dibincangkan dengan lebih terperinci di dalam bahagian yang akan datang.

Struktur kerangka sesuai digunakan bagi mewakili pengetahuan kewangan kerana ia dapat menyokong perwakilan pengetahuan berkenaan dengan konsep-konsep atau kejadian-kejadian yang stereotaip. Kerangka-kerangka yang mewakili subdomain-subdomain kewangan boleh diorganisasikan ke dalam bentuk taksonomi-taksonomi ("taxonomies") mengikut dua jenis perhubungan di antara mereka. Perhubungan yang dimaksudkan ialah :-

- (1) pautan keahlian ("member links"), yang mewakili keahlian kelas ("class membership"); dan
- (2) pautan subkelas ("subclass links"), yang mewakili kekandungan ("containment") sesuatu kelas, atau pengkhususan ("specialisation").

Pautan-pautan tersebut menunjukkan perhubungan di antara kerangka-kerangka dan sesuai digunakan untuk mewakili konsep warisan ("inheritance concept"). Konsep ini penting kerana ia membenarkan kerangka peringkat bawah (iaitu kerangka anak) mewarisi sifat-sifat kerangka peringkat atas (iaitu kerangka bapa). Oleh demikian, tiada pembaziran akan timbul dari segi penyimpanan maklumat. Tambah pula, setiap slot dalam kerangka boleh diisi oleh struktur kerangka yang lain, dan sesuatu kerangka boleh mengisi beberapa slot berlainan di dalam kerangka-kerangka yang berlainan.

Formalisa kerangka membenarkan suatu mekanisma pengisian slot dan instantasi kerangka. Apabila sesuatu kerangka diinstantasikan, pentaabiran ("inferences") dapat dilaksanakan yang berkaitan dengan keadaan sebenarnya. Sebagai contoh, di dalam kerangka **depreclatlon**, apabila proses susutnilai telah dikenalpastikan, kesimpulan-kesimpulan berkenaan dengan aset tetap ("fixed assets") yang disusutnilaikan ("depreciated"), kaedah-kaedah susutnilai yang digunakan dan sebagainya boleh ditaabirkan. Sebahagian daripada pentaabiran ("inferences") ini mungkin terkandung secara langsung di dalam kerangka itu sendiri (melalui slot - slot lalai ("default")); pentaabiran lainnya mungkin dilaksanakan mengikut fakta yang mengatakan bahawa pada kebiasaannya sesuatu kerangka itu berhubungan dengan kerangka-kerangka lain melalui pautan-pautan yang telah dinyatakan tadi. Konsep ini dinamakan sebagai konsep warisan.

Jika dibandingkan dengan jenis perwakilan pengetahuan yang lain, perwakilan ini adalah lebih baik kerana ia membenarkan penaakulan yang cekap, terutama sekali melalui penggunaan nilai-nilai lalai dan konsep warisan. Perwakilan kerangka juga membenarkan proses penaakulan dilaksanakan, walaupun kadangkala fakta-fakta yang ada tidak terdapat dengan secara eksplisit atau menyeluruh. Ciri ini amat berguna bagi domain masalah kami, kerana proses penaakulan perlu dilaksanakan memandangkan maklumat yang diperolehi biasanya tidak lengkap.

Sifat lain yang penting ialah penggunaan tatacara atau prosedur ("procedural attachment"), dengan kemungkinan maklumat prosedur dikaitkan kepada slot-slot tertentu. Prosedur-prosedur ini ditetapkan secara eksplisit dan ia diperlukan untuk membuat pelbagai kesimpulan tentang pengisian slot atau untuk memastikan kesahan keadaan selepas slot-slot diisikan. Ciri ini adalah penting untuk memerihalkan langkah-langkah pengiraan yang perlu digunakan di dalam kaedah-kaedah susutnilai yang berlainan, terutama sekali apabila sesuatu slot ingin diisikan.

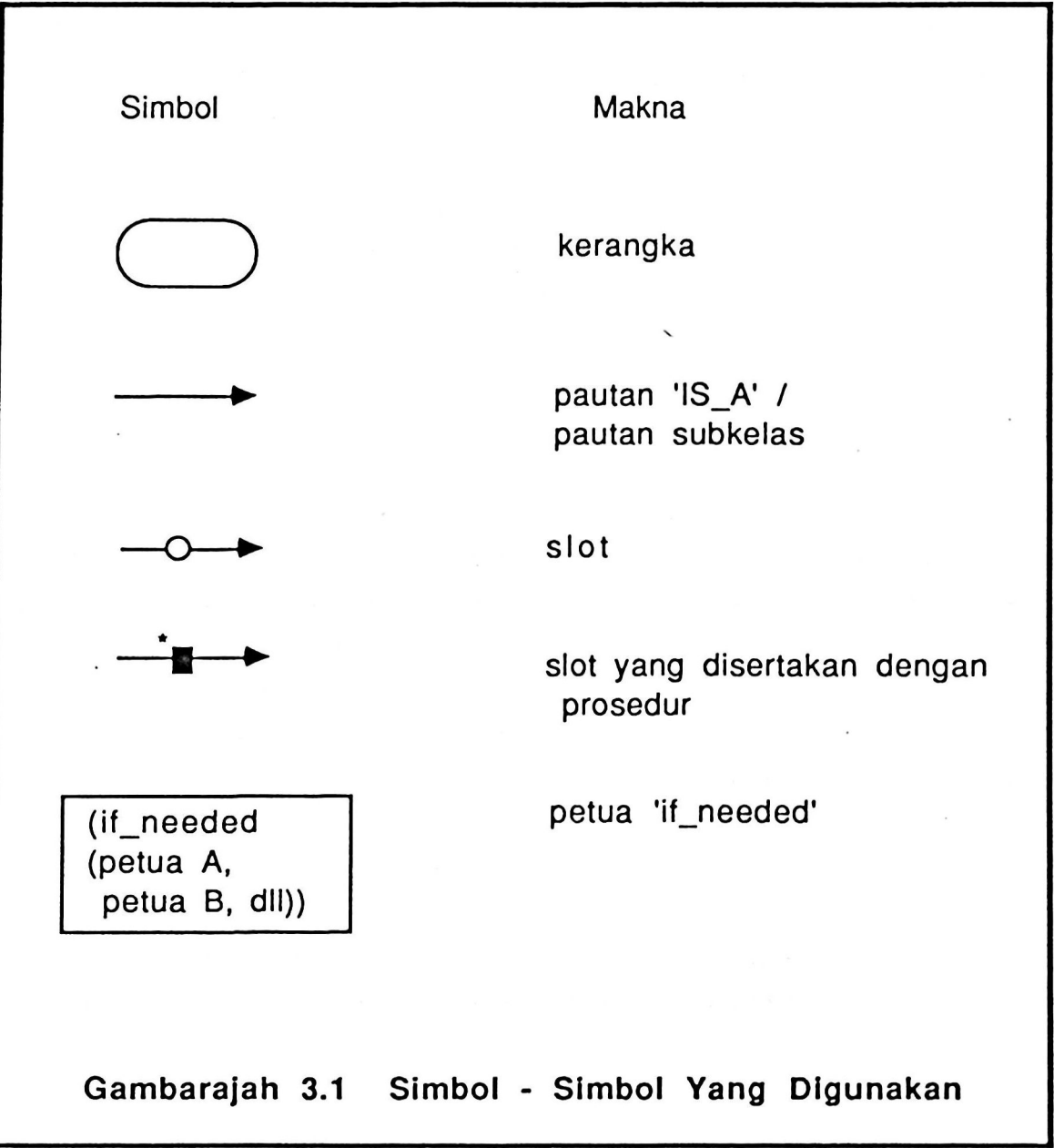
[Lim 89] juga menyenaraikan beberapa kebaikan terhadap kegunaan kerangka untuk mewakili pengetahuan kewangan iaitu :-

- (1) ia memberi pandangan atau gambaran dari segi cara pemikiran manusia terhadap penyata-penyata kewangan dan nota-nota kaki;

- (2) ia menyediakan suatu perwakilan yang ringkas dan tepat bagi penghuraian hubungan-hubungan yang berguna;
- (3) ia menyokong kaedah pemerihalan-melalui-pengkhususan ("definition- by-specialisation") yang ringkas dan tepat;
- (4) ia berupaya menyokong aktiviti-aktiviti penaakulan ("reasoning activities").

3.3 Tatatanda Dan Simbol Yang Digunakan

Untuk memudahkan perbincangan, gambarajah-gambarajah yang memperlihatkan rekabentuk kerangka akan digunakan. Oleh itu, sebelum rekabentuk tersebut dapat diperbincangkan dengan terperinci, kami memperkenalkan dahulu tatatanda-tatatanda dan simbol-simbol yang diperlukan. Ternyata di sini bahawa tatatanda-tatatanda dan simbol-simbol ini boleh juga digunakan untuk domain aplikasi lain yang menggunakan struktur kerangka. Gambarajah 3.1 meringkaskan simbol-simbol yang dimaksudkan.



Di sini, elok juga kami memperkenalkan nilai-nilai slot yang digunakan, iaitu :-

(1) nilai asas umum seperti integer, nyata, boolean, rentetan, dsb.

(2) senarai nilai asas ("list of values"), yang ditunjukkan dengan kurungan segiempat ([]); pada sesuatu ketika, hanya satu daripada nilai asas di dalam senarai ini boleh diperuntukkan kepada slot;

(3) set nilai asas ("set of values"), yang ditunjukkan dengan rangkaikata "set of"; pada sesuatu ketika, suatu subset, iaitu satu nilai atau lebih, boleh diperuntukkan kepada slot;

(4) **method**, yang biasanya disertakan dengan nama method atau prosedur itu; suatu prosedur menjanakan nilai asas yang diperuntukkan kepada slot.

(5) nilai kekangan yang ditunjukkan di dalam kurungan kerinting ({ }); sesuatu nilai slot itu perlu mengikuti syarat-syarat yang dikenakan oleh suatu kekangan.

(6) nilai lalai yang ditunjukkan di dalam kurungan biasa.

Perhatian: (5) dan (6) tidak digunakan di dalam rekabentuk kami kerana tiada nilai lalai dan kekangan yang sesuai dapat ditetapkan. Bagaimanapun kemungkinan ia akan digunakan di dalam penyelidikan masa depan atau untuk domain yang lain sekiranya didapati sesuai.

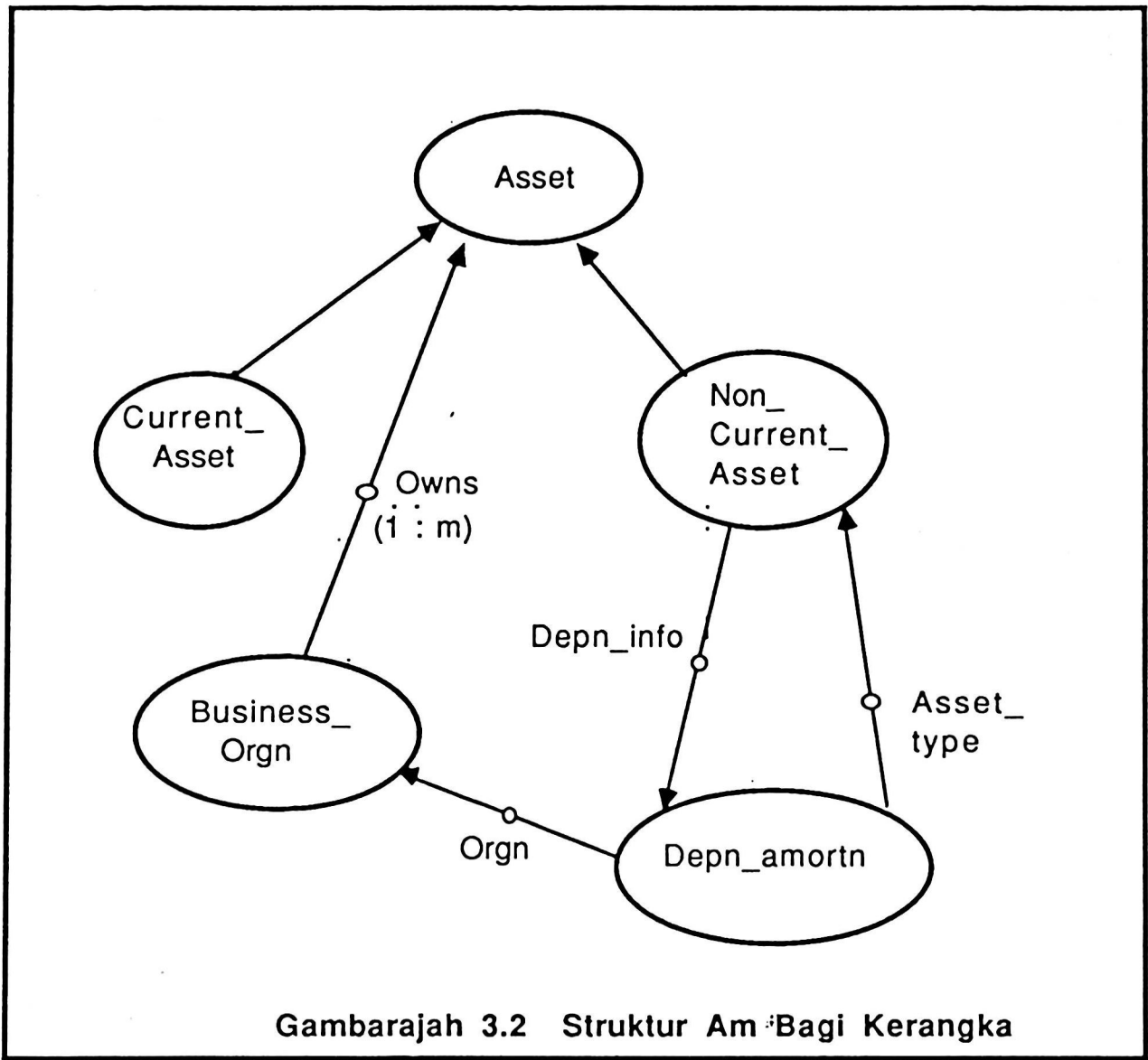
Bagi petua "if_needed" yang mungkin juga dikaitkan kepada sesuatu slot, simbol segiempat tepat yang mengandungi petua-petua berkaitan digunakan (seperti yang ditunjukkan dalam gambarajah 3.1). Petua-petua ini akan ditulis di dalam format tertentu, iaitu seperti diperlihatkan di bawah:-

nilai1 : petua1,
nilai2 : petua2,
dan lain - lain lagi.

Bagi sesuatu baris, nilai di sebelah kiri mestilah benar sekiranya petua di sebelah kanan itu sah. Jika petua1 tidak sah maka petua2 pula akan diuji, dan begitu juga seterusnya. Nilai yang pertama diperolehi diperuntukkan kepada slot. Petua "if_needed" ini hanya dilaksanakan apabila diperlukan. Tatatanda "a --> b" yang bermaksud slot b di dalam kerangka a juga digunakan di dalam ungkapan "X : a --> b" dengan tafsiran "X ialah nilai slot b bagi kerangka a (kemungkinan a juga merupakan ketikaan (instance) bagi sesuatu kerangka yang lain), ataupun ungkapan "Y : X : a --> b" yang bermaksud "nilai Y adalah benar dan diperuntukkan kepada sesuatu slot jika dan hanya jika petua "X : a --> b" adalah sah".

3.4 Rekabentuk Struktur Kerangka

Secara amnya, rekabentuk kerangka yang kami gunakan (iaitu khusus bagi domain susutnilai) mempunyai tiga kelas yang utama, iaitu **asset**, **depreciation** dan **business_orgn**, sebagaimana yang ditunjukkan di dalam gambarajah 3.2.



Gambarajah 3.2 Struktur Am 'Bagi Kerangka

Kelas-kelas ini diwakili oleh kerangka **asset** (yang membentuk suatu hierarki dengan kerangka **current_asset** dan **non_current_asset**), kerangka **depn_amortn** (singkatan bagi "depreciation/amortisation") dan kerangka **business_orgn**. Kedua-kedua kerangka **depn_amortn** dan **business_orgn** wujud secara tersendiri, iaitu tidak mempunyai pautan subkelas atau pautan 'IS_A'.

3.4.1 Kerangka Asset

Menurut [Wilcox 80], aset ("asset") ialah sumber-sumber yang dipunyai sesuatu perniagaan atau perkara-perkara yang diberi hak kepada sesuatu perniagaan untuk digunakan. Aset menyumbangkan keuntungan masa depan kepada sesuatu firma melalui jualan atau penggunaannya. Mengikut pengetahuan kewangan, aset dapat dikategorikan kepada dua subkelas yang utama, iaitu aset semasa ("current asset") dan aset tak semasa ("non current asset"). Kedua-dua subkelas ini mempunyai ciri-ciri yang berlainan seperti yang akan diperbincangkan di sini.

[Wilcox 80] menakrifkan aset semasa sebagai wang tunai, yang boleh luput ataupun sedia digunakan oleh operasi-operasi tertentu dalam tempoh setahun atau dalam tempoh pusingan operasi perniagaan biasa. Aset bukan semasa pula merupakan aset yang mempunyai jangka hayat yang panjang dan digunakan untuk mendapatkan keuntungan semasa operasi-operasi biasa sesuatu perniagaan dilaksanakan dan juga untuk

menyumbangkan keuntungan dalam masa lebih daripada satu tahun. Berasaskan konsep-konsep ini, rekabentuk kerangka kami menetapkan aset sebagai suatu kerangka yang dihubungkan (oleh pautan 'IS_A') kepada dua kerangka anak iaitu kerangka `current_asset` dan `non_current_asset` sebagaimana yang ditunjukkan di dalam gambarajah 3.2.

Oleh kerana sesuatu aset itu mempunyai jangka hayat, kerangka `aset` diberikan slot `consumption_perlod` (bernilai integer) yang menunjukkan jangka masa (dalam tempoh tahun) sesuatu aset itu boleh digunakan dalam sesuatu operasi perniagaan (sebagaimana yang ditunjukkan di dalam gambarajah 3.3). Maklumat jangka masa ini juga penting khusus bagi domain susut nilai oleh kerana kemungkinan ia diperlukan untuk pengiraan nilai susut aset itu (suatu perkara yang akan dibincangkan di bahagian 3.3.2 nanti).

Kerangka `non_current_asset` mempunyai slot `type_of_asset`, `depn_info` dan `acquisition_cost`. Slot `type_of_asset` mengkategorikan aset-aset bukan semasa seperti yang dinyatakan di dalam [Wilcox 80], khususnya :-

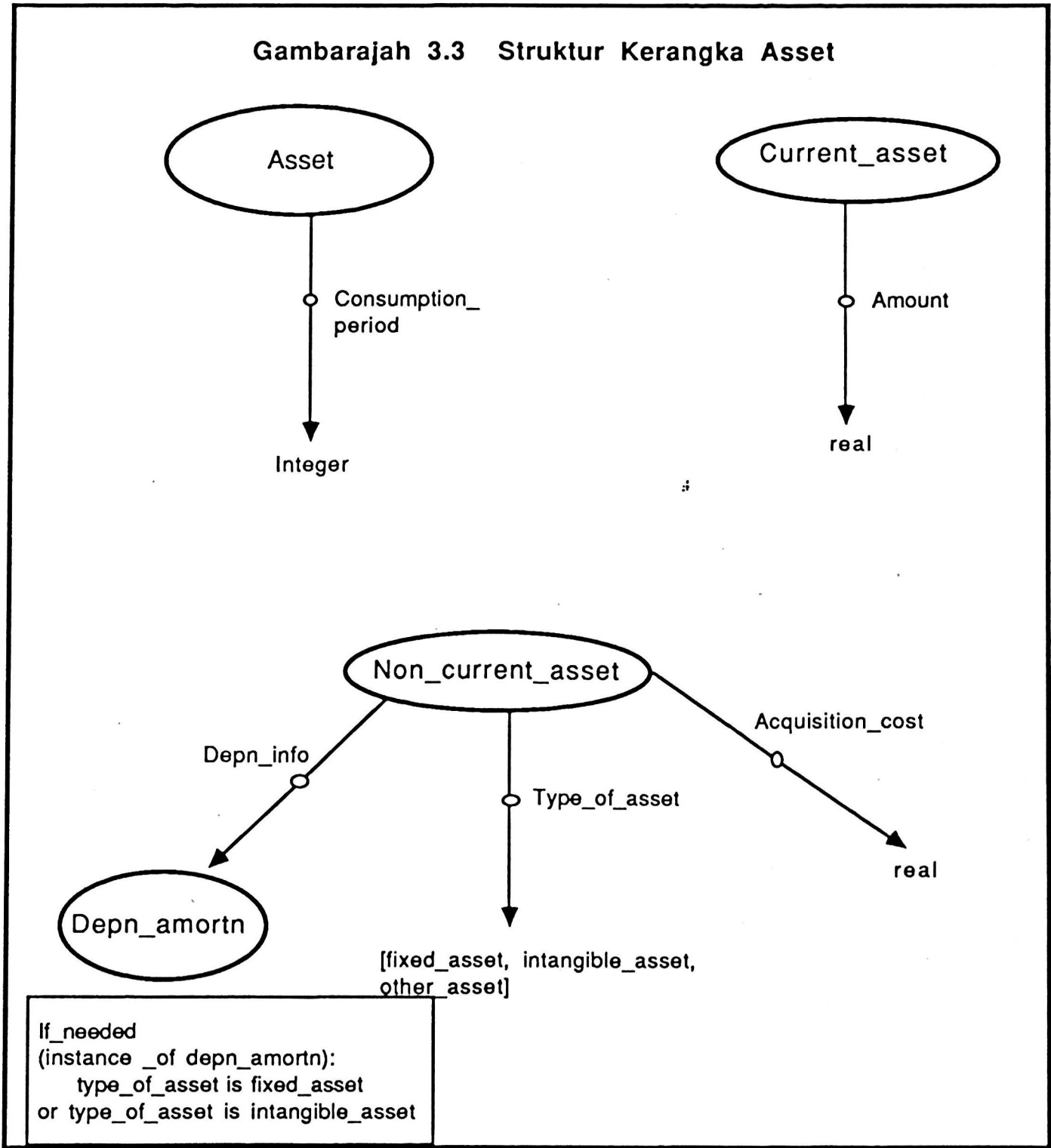
- (1) aset tetap ("fixed asset");
- (2) aset tak ketara ("intangible asset");
- (3) aset lain ("other asset").

Oleh itu, slot tersebut mempunyai senarai nilai `fixed_asset`, `intangible_asset` dan `other_asset` (seperti yang diperlihatkan di dalam gambarajah 3.3). Aset tetap terdiri daripada harta-benda ("properties") yang wujud secara fizikal (seperti tanah, bangunan, alat kelengkapan dan sebagainya) yang boleh digunakan di dalam operasi-operasi biasa sesuatu perniagaan. Aset tetap yang tidak boleh digunakan di dalam operasi perniagaan biasa, seperti aset-aset di dalam keadaan terbiar ("idle"), dalam penyelenggaraan ("maintenance"), dalam instalasi atau penempatan ("under installation") dan yang boleh dipulangkan ("returnable"), semuanya termasuk di dalam kategori aset lain ("other assets"). Aset tak ketara ("intangible asset") pula merupakan aset bukan fizikal (sebagai contoh, trademark, patent, hakcipta dan franchise). Menurut [Wilcox 80] lagi, harta-benda dan peningkatan ("improvements") yang dipajak ("leased") dikategorikan sebagai aset tak ketara, walaupun aset-aset ini wujud secara fizikal. Ini berdasarkan hakikat bahawa syarikat yang memajak aset tersebut bukannya pemilik, tetapi hanya merupakan pemilik hak penggunaannya.

Slot `depn_info` pula merakamkan kewujudan proses-proses susut nilai ("depreciation") dan pelunasan ("amortisation") yang dikaitkan kepada aset tak semasa. Pada amnya, suatu aset tak semasa berjenis tetap (mengikut slot `type_of_asset`), disusutnilaikan ("depreciated") walhal aset tak ketara ("intangible asset") dipelunaskan ("amortised"). Aset lain ("other asset") langsung tidak melalui proses-proses ini. Aset tetap yang mempunyai hayat yang panjang (khususnya tanah yang tidak dipajak) juga tidak disusutnilaikan. Berdasarkan maklumat ini, slot `depn_info` diinstantasikan dengan kerangka `depn_amortn` sekiranya sesuatu aset itu berkategori tetap dan berhayat pendek ataupun apabila ia berkategori tak ketara. Sekiranya aset itu berkategori lain ("other aset"), slot `depn_info` langsung tidak akan diinstantasikan .

Slot ketiga dalam kerangka `non_current_asset`, iaitu slot `acquisition_cost` (yang mempunyai nilai nyata) merakamkan harga perolehan aset itu. Harga ini diperlukan untuk pengiraan nilai susut, suatu perkara yang akan dibincangkan di bahagian 3.3.2.

Perhatian tidak akan ditumpukan kepada kerangka `current_asset` kerana (mengikut [Wilcox 80]) aset semasa langsung tidak mengalami susutnilai, dan dengan itu tidak termasuk di dalam tumpuan penyelidikan. Bagaimanapun, suatu slot mungkin diperlukan di sini, iaitu slot `amount` yang menyimpan jumlah nilai aset yang bernilai nyata.

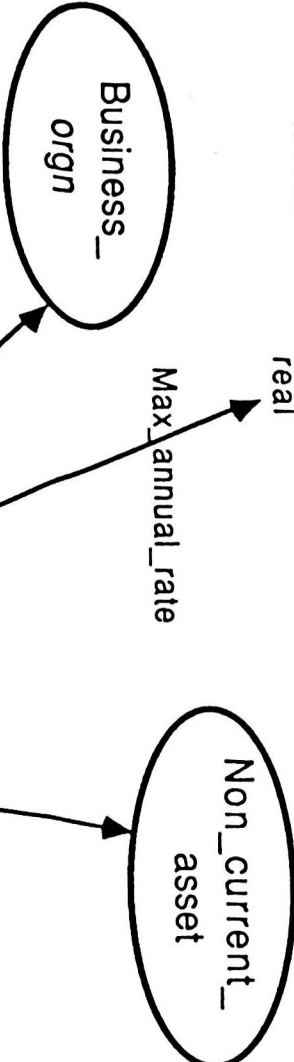


am araja . ru tur erang a epn_amortn

```

(if_needed
(R: (P : asset_type->
consumption_period),
R is 1/P ))

```



```

method :
mstraight_line (V, R) <=
Cost : (asset_type -> acquisition_cost),
S : ask_user(What is the salvage value?),
V is R * (Cost - S).
mreducing_balance (V, R) <=
Cost : (asset_type -> acquisition_cost),
T : ask_user(What is the total depreciation
of previous periods?),
V is R * (Cost - T).

```

```

(if_needed
(depreciation:
fixed_asset: asset_type ->
type_of_asset),
(amortisation:
intangible_asset:
asset_type ->
type_of_asset))

```

```

(if_needed
(R: (P: asset_type ->
consumption_period),
R is 1/P ) )

```

```

(if_needed
(mstraight_line (V, max_annual_rate):
method_used is atraight_line),
(mreducing_balance(V,max_annual_rate)
: method_used is reducing_balance))

```

```

(if_needed
(mstraight_line (V, min_annual_rate):
method_used is straight_line)
set of
string
(mreducing_balance (V, min_annual_rate):
method_used is reducing_balance))

```

```

(if_needed
([straight_line &reducing_balance] :
process is depreciation),
([straight_line]: process is amortisation))

```

3.4.2 Kerangka Depn_amortn

Seperti yang dicadangkan terlebih awal, kerangka **depn_amortn** mewakili proses susutnilai ("depreciation") dan pelunasan ("amortisation"). Proses-proses ini diwakili oleh hanya satu kerangka kerana ciri-cirinya adalah sama [Wilcox 80], iaitu proses peruntukkan kos aset terhadap hayatnya agar perbandingan dapat dibuat di antara kos dengan keuntungan terhasil. Beza yang terdapat di antara dua proses ini hanyalah dari segi seperti yang disebutkan terlebih awal, iaitu proses susutnilai dikhususkan kepada aset tetap manakala proses pelunasan dikhususkan kepada aset tak ketara.

Kerangka ini mempunyai lapan slot, iaitu **asset_type**, **process**, **method_used**, **min_depn_value**, **max_depn_value**, **max_annual_rate**, **min_annual_rate** dan **orgn** (lihat gambarajah 3.4), yang diuraikan berikutnya.

3.4.2.1 Slot Asset_type Dan Slot Process

Slot **asset_type** dikaitkan kepada kerangka **non_current_asset**, kerana sebagaimana yang telah dinyatakan terlebih awal, hanya aset tak semasa sahaja dibenarkan melalui proses susutnilai atau pelunasan. Slot ini wujud untuk menunjukkan aset yang disusutnilai atau dipelunaskan.

Slot **process** juga diperlukan untuk menandakan sama ada proses bagi **depn_amortn** merupakan proses susutnilai atau pelunasan. Sebagaimana yang ditunjukkan dalam gambarajah 3.4, slot ini mempunyai nilai jenis senarai rentetan yang mengandungi **depreciation** dan **amortisation**. Pada sesuatu ketika, slot ini boleh mempunyai hanya satu nilai. Petua "if_needed" disertakan untuk memastikan bahawa slot itu mengambil nilai **depreciation** jika **asset_type** (iaitu yang mewakili ketikaan bagi kerangka **non_current_asset**) mempunyai slot **type_of_asset** yang bernilai **fixed_asset** (sila rujuk kepada bahagian 3.3.1), manakala nilai **amortisation** diambil sekiranya slot **type_of_asset** bernilai **Intangible_asset**.

3.4.2.2 Slot Method_used

Slot **method_used** yang merakamkan kaedah susutnilai yang digunakan dihubungkan dengan set nilai rentetan (iaitu nama semua kaedah yang mungkin digunakan). Mengikut [Wood 67], kaedah-kaedah susutnilai ini terdiri daripada kaedah garis_lurus ("straight_line") dan kaedah baki_berkurangan ("reducing_balance"). Walaupun kaedah-kaedah lain mungkin sesuai digunakan bagi kes-kes tertentu, kaedah garis_lurus merupakan yang paling baik, diikuti kaedah baki_berkurangan. Rekabentuk kami menekankan dua jenis kaedah utama ini, dan mengabaikan kaedah-kaedah lain. Petua "if_needed" juga disertakan untuk memastikan yang slot itu diisikan dengan nilai-nilai **straight_line** dan juga **reducing_balance**, jika slot **process** bernilai **depreciation**, walhal hanya nilai **straight_line** diisikan jika slot **process** bernilai **amortisation**. Pilihan ini berasaskan hujah [Wilcox 80] yang mengatakan bahawa perbezaan di antara proses susutnilai dan pelunasan ialah yang berikut:-

- (1) susutnilai melibatkan aset tetap, manakala pelunasan melibatkan aset tak ketara;
- (2) pelunasan sesuatu aset tak ketara biasanya menggunakan kaedah garis_lurus.

Jenis set nilai dihubungkan dengan slot ini kerana pada kebiasaannya sesuatu organisasi dan subsidiari-subsidiarinya menggunakan lebih daripada satu kaedah susutnilai atau pelunasan.

3.4.2.3 Slot **Min_annual_rate**, **Max_annual_rate**, **Slot Min_depn_value**, **Max_depn_value** Slot **Orgn**

Slot **min_annual_rate** dan **max_annual_rate** (yang mempunyai nilai nyata) mewakili kadar maksimum dan kadar minimum bagi pengiraan sesuatu nilai susut atau pelunasan. Sebagaimana yang ditunjukkan dalam petua "if_needed" di dalam gambarajah 3.4, rumus pengiraan kadar-kadar tersebut berbentuk berikut:-

$$\text{Kadar} = 1 / \text{hayat_aset}$$

yang mengatakan bahawa kadar itu diperolehi dengan membahagikan nilai 1 dengan nilai slot **consumption_period** dalam **asset_type**.

Slot **min_depn_value** dan **max_depn_value** menyimpan nilai-nilai susutnilai atau pelunasan yang terhasil dari kaedah atau prosedur yang disertakan. Pengiraan nilai-nilai ini bergantung kepada kadar susutnilai atau pelunasan serta kaedah-kaedah susutnilai yang diberikan. Bagi kaedah **garis_lurus** dan **baki_berkurangan**, rumus-rumus berkaitan adalah seperti yang diberikan oleh [Wood 67], iaitu :-

bagi **garis_lurus**,

$$(\text{kos} - \text{nilai_sisaan}) / \text{hayat_aset}$$

bagi **baki_berkurangan**,

$$\text{kadar} * (\text{kos} - T), \text{ di mana } T \text{ ialah hasil tambah nilai susut bagi tempoh masa yang lalu.}$$

Petua "if_needed" menentukan kaedah yang perlu digunakan untuk mengirakan nilai-nilai susut itu. Seperti yang ditunjukkan dalam gambarajah 3.4, bagi kaedah **straight_line**, nilai sisaan diperolehi daripada pengguna secara langsung semasa pengiraan dilaksanakan, dan begitu juga bagi kaedah **reducing_balance** untuk hasil tambah T. Nilai-nilai tersebut tidak ditetapkan sebagai slot kerana penggunaannya terhad kepada pengiraan nilai susut sahaja, dan juga hanya untuk kaedah-kaedah tertentu sahaja.

Slot **orgn** pula dihubungkan dengan kerangka **business_orgn** untuk mewakili organisasi yang terlibat dengan proses susutnilai atau pelunasan tersebut.

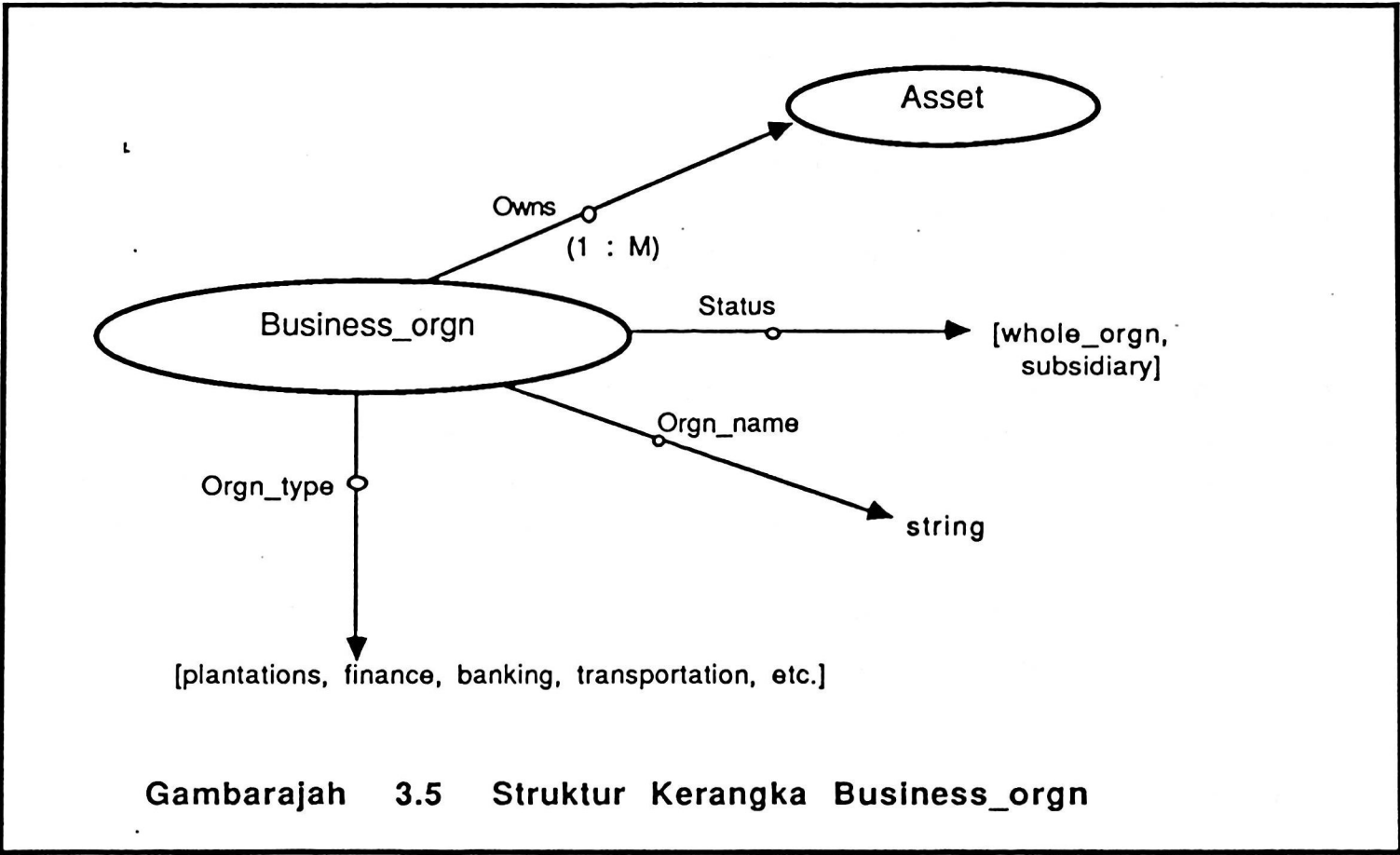
3.4.3 Kerangka **Business_orgn**

Kerangka **business_orgn** mewakili organisasi perniagaan yang terlibat dalam domain masalah kami. Seperti yang ditunjukkan dalam gambarajah 3.5, ia mempunyai empat slot iaitu **owns**, **orgn_name**, **status** dan **orgn_type**.

Slot **owns** merakamkan aset-aset yang dipunyai oleh organisasi itu. Ia dihubungkan dengan kerangka **aset** melalui perhubungan satu-ke-banyak (1:m) kerana pada amnya sesebuah organisasi itu mempunyai lebih daripada satu aset.

Slot **status** pula mewakili status organisasi itu iaitu samada ketikaan organisasi itu merupakan keseluruhan organisasi ataupun hanya subsidiari sesuatu organisasi sahaja. Keseluruhan organisasi bermaksud cawangan utama bersama semua subsidiarinya. Nilai-nilai yang terlibat bagi slot ini ialah **whole_orgn** dan **subsidiary**, seperti yang ditunjukkan di dalam gambarajah 3.5.

Slot **orgn_name** menamakan organisasi tersebut, manakala slot **orgn_type** merakamkan jenis organisasi itu, seperti pertanian, kewangan, dan sebagainya.



3.5 Kesimpulan Bagi Rekabentuk

Rekabentuk struktur kerangka yang kami bentangkan hanya merupakan satu contoh yang tidak begitu menyeluruh. Bagaimanapun, kami percaya bahawa contoh ini bersifat pedagogi yang baik dan boleh diperkembangkan untuk menerapkan domain aplikasi lain.

Walaupun rekabentuk struktur kerangka kami agak ringkas, ia boleh dikatakan cekap dan dapat mewakili pengetahuan masalah domain susutnilai secara menyeluruh. Rekabentuk yang kompleks dari segi perwakilan tidak semestinya baik dan cekap kerana kompleksitinya mungkin disebabkan perwakilan maklumat yang berlebihan ("redundancy") yang seterusnya disingkirkan, ataupun disebabkan perwakilan-perwakilan yang tidak relevan dengan domain pengetahuan tumpuan. Kami juga cuba mengelak penggunaan slot berjenis boolean kerana nilai yang boleh diisi bagi slot jenis ini hanyalah *true* atau *false* sahaja, yang kami dapati kurang memuaskan untuk perwakilan pengetahuan.

3.6 Perwakilan Struktur-M

Tujuan muktamad kami adalah untuk menganalisis nota-nota kaki ini dan menyimpan maklumat yang diperolehi di dalam bentuk kerangka. Bagaimanapun, untuk memastikan bahawa analisis ini dijalankan dengan sempurna, suatu analisis bahasa asli yang menggunakan kaedah linguistik berkomputer diperlukan. Suatu kaedah analisis seperti ini sudah sedia dilaksanakan dalam bidang penterjemahan berkomputer. Satu daripadanya menggunakan struktur-m sebagai struktur perwakilan.

Di dalam penyelidikan kami, input merupakan nota kaki di dalam bentuk bahasa asli (khusus bagi domain susutnilai). Struktur-m dipilih sebagai struktur perwakilan kerana ia direkabentuk khusus untuk memperwakilkan empat paras tafsiran [Vauquois 78, Zaharin 87], iaitu tafsiran permukaan, fungsi, logik dan semantik. Tafsiran-tafsiran ini sama dengan yang telah dibincangkan di dalam bab 2.

Perwakilan struktur-m berbentuk struktur pohon beranotasi ("annotated tress structure"). Untuk mewakili maklumat-maklumat yang disebutkan di atas, setiap nod ("node") di dalam pohon perwakilan dianotasikan dengan satu set ciri, dengan kemungkinan setiap ciri mempunyai struktur dalaman yang kompleks, umpamanya di dalam bentuk pohon berlabel tidak bertertib ("unordered labeled tree"). Khususnya, setiap nod diberikan sekurang-kurangnya tiga ciri :-

(1) kelas sintagmatik ("syntagmatic class") [seperti NP ("Noun Phrase"), VK ("Verbal Kernel"), PP ("Prepositional Phrase"), dan sebagainya]; atau kelas sintaksis ("syntactic class") [seperti N ("Noun"), V ("Verb"), DET ("Determiner"), dan sebagainya].

(2) fungsi sintaksis ("syntactic function") [seperti SUBJ ("Subject"), GOV ("Governor"), DES ("Designator"), dan sebagainya];

(3) hubungan logik ("logical relation") [seperti ARG0 ("argument 0"), ARG1 ("argument 1") dan sebagainya]; atau hubungan semantik ("semantic relation") seperti INST ("Instrument"), LOC ("Location"), METHOD, dan sebagainya].

Pada amnya, akar pohon mempunyai hanya kelas sintagmatik manakala pada daun-daun pohon ("leaves"), hubungan logik merupakan unit leksikon ("lexical unit") yang berpadanan dengan perkataan di dalam ayat ataupun sesuatu hubungan semantik (dengan unit leksikon diberikan di dalam kurungan) sekiranya unit leksikon itu tidak memainkan peranan yang penting di dalam maksud ayat.

Pemetaan di antara perwakilan struktur-m dengan perwakilan kerangka boleh direkabentuk melalui tafsiran-tafsiran yang ada, iaitu permukaan, fungsi, logik dan semantik. Pemetaan ini akan dibincangkan secara mendalam di dalam bab 4. Buat masa kini, kita lihat kepada struktur-m dengan lebih terperinci, khususnya paras tafsiran logik.

3.7 Struktur-M dan Peranan Semantik

Seperti yang disebutkan terlebih awal, struktur-m memberikan empat paras tafsiran iaitu permukaan, fungsi, logik dan semantik. Apendiks B dalam disertasi ini memperlihatkan perwakilan struktur-m bagi ayat-ayat di dalam korpus.

Sebelum strategi pemetaan dapat direkabentuk, peranan-peranan PRED, PATIENT dan BENEf (yang merupakan perhubungan semantik) dalam teks input perlu dikenalpasti.

Kami juga memperkenalkan peranan-peranan seperti METHOD dan DOMAIN_CALC untuk memberi tafsiran semantik bagi domain kewangan, khususnya susutnilai. METHOD memberi tafsiran kepada kaedah-kaedah susutnilai atau pelunasan yang digunakan bagi menghitung nilai susut suatu aset (sebagaimana yang telah diperbincangkan dalam bahagian 3.4), walhal DOMAIN_CALC pula memberi tafsiran kepada jangkamasa ("period") yang digunakan untuk pengiraan nilai susut suatu aset itu. Lain peranan tidak perlu diperbincangkan di sini, kerana ini akan menyebabkan perbincangan yang panjang tentang linguistik, yang langsung tidak berkaitan dengan tugas penyelidikan kami. Kami harap nama-nama yang digunakan khusus untuk peranan-peranan itu dapat menjelaskan maksudnya secara tersendiri.

Peranan-peranan logik yang digunakan dalam perwakilan struktur-m bagi ayat korpus kami ialah ARG0, ARG1, ARG2, GRAO (terbalik bagi ARG0), dan sebagainya. Seperti yang dibincangkan di dalam bab 2, ARG0 merupakan AGENT, ARG1 merupakan PATIENT dan ARG2 pula merupakan BENEf.

Jelas dari perbincangan yang lalu, perhubungan-perhubungan semantik [seperti PATIENT, AGENT, BENEf, dan sebagainya] dan logik [seperti ARG0, ARG1, ARG2, dan sebagainya] berhubungan dengan prediket utama (PRED), yang merupakan proses. Tambahan pula, perhubungan-perhubungan semantik dan logik ini diperihalkan secara unggul bagi setiap proses atau PRED. Sesuatu PRED boleh dipandang sebagai satu templet seperti berikut:-

PRED (AGENT, PATIENT, INST, BENEf, METHOD,)

ataupun

PRED (ARG0, ARG1, INST, ARG2, METHOD, DOMAIN_CALC,)

di mana ayat sebenar akan membentuk instantasi kepada salah satu daripada templet tersebut.