

Artigos

Dialogos

ANTONIO GOUVEIA PORTELA

Professor Catedrático
Departamento de Sistemas I.S.T.
Lisboa,

Julho de 1987

RESUMO

Apresentação de uma linguagem (F_LING) e de um dialogante formal (F_DIALOG) que interacciona com o "mundo exterior" representado por um dialogante "humano" (H_DIALOG) que usa uma linguagem natural (N_LING).

Em (F_LING) introduz-se o conceito de "frases holónicas" (HOL) valorizadas quanto a respectiva credibilidade.

Nos capítulos finais discutem-se os temas: tradução (nos dois sentidos), inferência e indução.

RESUME

Presentation du langage (F_LING) et du dialoguant formel (F_DIALOG) lequel interagit avec le "monde extérieur" représenté par le "dialoguant humain" (H_DIALOG) qui emploie un langage naturel (N_LING).

On introduit en (F_LING) le concept de "frases holoniques" (HOL) valorisées quant à leurs respectives crédibilités.

Les derniers chapitres sont consacrés aux thèmes suivants: traduction (dans les deux sens), inférence et induction.

ABSTRACT

Presentation of the language (F_LING) and the "formal dialoguer" (F_DIALOG) who interacts with the "outside world" represented by a "human dialoguer" (H_DIALOG) who employs natural language (N_LING).

The concept of "holonic sentences" (HOL) is introduced in (F_LING), and "credibility values" are applied to these sentences.

The following subjects are discussed in the last chapters: translation (version and retroversion), inference and induction.

A) Introdução: Diálogo

As sociedades humanas criam ou adaptam uma Linguagem Natural (N_LING) que vai autorizar o diálogo e a comunicação entre os membros dessa sociedade.

As principais características duma linguagem (N_LING) são:

- dispor de regras sintáticas na formação de Frases a partir das Palavras.
- a interpretação das palavras e frases ser razoavelmente uniforme nessa sociedade.
- a ambiguidade duma comunicação poder ser reduzida se entre o emissor da mensagem e o receptor da mesma se estabelecer um diálogo (DIALOG), até completa satisfação das partes intervenientes.

Numa sociedade humana moderna, nenhum membro faz uso de todas as palavras postas à sua disposição, nem emprega todas as frases susceptíveis de construção. Assim, haverá muitos dialectos, entre os quais avultam os especializados, e.g.: científicos, jurídicos, poéticos, etc.

O objectivo desta comunicação consiste em apresentar um dialogante formal (F_DIALOG) com a capacidade de receber comunicações de humanos (H_DIALOG) e manter um diálogo com estes com o fim de esclarecer e precisar o sentido semântico das frases em linguagem natural (N_LING) e "traduzir" essas frases numa linguagem formal (F_LING).

Numa segunda fase, responder a perguntas propostas pelos dialogantes humanos (H_DIALOG) e dum modo geral inferir. A faculdade de indução será também possível em certas condições referidas no Cap. E.

As principais características da (F_LING) são:

- * Introdução do conceito de "frase holónica", que, dum modo singelo, se pode definir como uma frase não susceptível de ser 'construída' a partir de frases ainda mais simples. A frase holónica possui no máximo 3 palavras e a palavra central designa-se por "verbo". (Cap.: B e seguintes).
- ** O emprego de valorisantes (índices de pertença) para apreciar a "credibilidade" ou "verdade" das frases. Para este efeito usa-se a teoria dos conjuntos_vagos (fuzzy sets)
- *** A introdução de certas faculdades inatas com vista aos processos de indução.

Resumindo as principais funções do (F_DIALOG) são:

- a) Comunicar com vários humanos, recebendo destes mensagens sob a forma de frases suportadas numa linguagem natural de referência (N_LING) que é inteligível para o dialogante formal (F_DIALOG).
- b) Questionar o dialogante humano (H_DIALOG), até completa satisfação do (F_DIALOG).

c) Traduzir a mensagem humana já esclarecida numa linguagem formal (F_LING) que é um dialecto formal (F_DIALECT) da linguagem natural associada (N_LING).

d) Porque a linguagem (F_LING) é formal, será possível inferir com rigor, deduzindo frases escritas em (F_LING).

e) Dado que (F_LING) é um (F_DIALECT) da (N_LING) será possível verter em linguagem natural (N_LING) as frases escritas em (F_LING).

f) Porque o conceito de frase holonica permite reconstruir frases e (F_DIALOG) é dotado de algumas faculdades inatas, em princípio, será possível realizar induções.

Recorde-se que a Linguagem Jurídica é um dialecto da linguagem natural que tem por propósito expresso eliminar as ambiguidades próprias da linguagem natural associada.

Com um propósito semelhante são construídas as linguagens especializadas das várias ciências, culminando este esforço nas linguagens formais (lógicas e matemáticas).

O objectivo do (F_DIALOG) e respectiva linguagem (F_LING) é idêntico na intenção de reduzir a ambiguidade e visa também a resolução de problemas formais que lhe sejam postos e, em certas circunstâncias induzir.

Como nota final, desejamos aqui sublinhar que este texto tem por fim apresentar discursivamente uma linguagem e não a sua sintaxe ou validar e suportar as acções feitas em trabalhos de outros autores.

Daf não ser apresentada bibliografia, nem demonstrações formais, e usar-se uma linguagem natural na respectiva exposição.

O autor, tenciona apresentar oportunamente, não só a estrutura da linguagem como os programas que hoje executam as operações referidas neste texto.

Estes programas estão sendo escritos em PASCAL que não é uma linguagem especializada ou especialmente apropriada.

B) Linguagem (F_LING)

B1) Introdução

A linguagem (F_LING) é suportada por Palavras (PLV), sucessões contíguas e finitas de caracteres. Numa frase (FRS) escrita em (F_LING), as palavras estão separadas entre si por espaços. Todas as palavras (PLV) estão inscritas num Dicionário (DIC).

No dicionário (DIC) não há palavras iguais, mas podem existir sinónimos embora as palavras respectivas sejam diferentes sucessões de letras.

Se na linguagem natural (N_LING) associada existir uma só palavra para representar conceitos diferentes, então, em (F_LING) acrescentam-se algumas indicações suplementares (letras e ou símbolos) para as distinguir.

Novas palavras podem ser acrescentadas ao Dicionário e designa-se por Estado do Dicionário (DIC_STT) o conjunto de Palavras existentes num dado momento.

Uma Triada palavras (PLV1, PLV2, PLV3) constitui uma Frase Holonica (HOL) se satisfizer as seguintes condições:

1) as três palavras estão contidas no Dicionário (DIC).

2) a respectiva ordem na declaração (HOL) é estrita.

3) a palavra central, e.g. PLV2, pertence a um tipo que será designado por verbo (VRB)

A leitura da triada (HOL) pode efectuar-se de dois modos apenas:

(>>>) da esquerda para a direita,
isto é, PLV1 invoca PLV3 por
intermédio de PLV2.

(<<<) da direita para a esquerda,
isto é, PLV3 invoca PLV1 por
intermédio de PLV2*.

A leitura do tipo (>>>) pode ser convertida numa leitura do tipo (<<<), substituindo (HOL) pela frase: (HOL*) :: (PLV3, PLV2*, PLV1), onde PLV2* designa-se por verbo inverso de PLV2.

A cada (HOL) está associado um Declarante ou Autor (AUT), pessoa, aparelho, ou conjuntos formados por pessoas e ou aparelhos. A palavra que descreve o Autor (AUT) está contida no Dicionário (DIC).

A cada (HOL) está associada, biunivocamente, uma palavra (FRS) que será o seu indicativo e está igualmente contida no Dicionário (DIC).

B2) Relação de (F_LING) com a Língua Natural (N_LING) e Autoria;

As frases na Língua Natural (N_LING) são complexas, isto é, contêm eventualmente várias frases Holónicas. Simboliza-se a frase natural por (N_FRS) e estes símbolos estão contidos no Dicionário (DIC).

Uma frase natural (N_FRS) toma a forma normal seguinte:

(AUT, N_FRS, PLV1, ..., PLVn)

Qualquer das Frases Holónicas (HOL) derivadas de uma frase natural terá a forma normal seguinte:

(FRS, AUT, N_FRS, PLV1, PLV2, PLV3).

Note-se que os índices 1..n, representam apenas a posição da palavra na frase e que PLV é um símbolo mais especificado de "Palavra".

Tipos (TIP) são palavras que designam conjuntos de palavras que possuem certos atributos comuns, e.g.: verbos, entes, atributos, advérbios, pronomes, etc. Os tipos (TIP) estão incluídos em (DIC).

Os tipos (TIP) podem ser estruturados taxonomicamente. A cada palavra numa Frase Natural ou Holonica, está associado o respectivo Tipo (TIP).

Credibilidade ou Valor (VLR) é a Função da Verdade ou Credibilidade contida numa frase e que, em geral, varia de palavra para palavra nessa frase.

A cada palavra numa frase em (N_LING) ou em (F_LING) associa-se um valor colhido num reticulado de referência.

A exposição deste tema será feita escolhendo para reticulado o intervalo [0,1], e as respectivas connectivas (max, min).

A forma completa como uma Frase Natural é recordada (N_REC), depois de 'esclarecer' o respectivo sentido pelo dialogante formal (F_DIALOG), é a seguinte:

(AUT, N_FRS, PLV1, ..., PLVn)

(TIPa, TIPf, TIP1, ..., TIPn)

(VLRa, VLRF, VLR1, ..., VLRn)

A forma completa como uma Frase Holonica é recordada (REC) é:

(FRS, AUT, N_FRS, PLV1, PLV2, PLV3)

(TIPh, TIPa, TIPf, TIP1, TIP2, TIP3)

(VLRh, VLRa, VLRF, VLR1, VLR2, VLR3)

B3) Património Informativo

O discurso Natural faz-se por meio de frases naturais (N_FRS) da responsabilidade de diversos autores.

Destas frases naturais são extraídas as frases holónicas (HOL) respectivas.

O património informativo (PTR_INF) é constituído pela reunião das frases Holónicas (HOL) e tal como sucede com o Dicionário (DIC), varia de conteúdo (ou património) e passa por diversos estados descritos, em cada momento, por (PTR_STT).

Recorda-se que as frases interrogativas têm cabimento e são encerradas como declarações hipotéticas que esclarecidas resolvem eventualmente um dado problema.

O (F_DIALOG), para induzir, deverá ser dotado de um conjunto de conceitos uns inatos outros adquiridos (Cap. E).

O contacto com o utilizador faz-se nos dois sentidos e em Linguagem Natural.

Foi criado um Conjunto de rotinas destinado a converter as frases naturais (N_FRS) em frases holónicas (HOL).

Essas rotinas permitem o diálogo com o dialogante humano e eliminar ou reduzir indefinições de expressão próprias da linguagem natural.

A este conjunto de rotinas dá-se o nome de "TRADUTOR".

Constituindo algum património informativo (PTR_INF) é possível extrair dele informação, não só na forma em como está arquivada como sob formas elaboradas, isto é, formas que resultam do processamento da forma (bruta) como foi e está memorizada.

B4) Conjuntos e Múltiplos

As principais operações consistem em formar conjuntos de Palavras, que representam quer entes, atributos, verbos, declarações ou outros conjuntos todos simbolizados por palavras contidas no dicionário (DIC).

Em geral um conjunto diz-se desordenado quando a ordem pela qual as palavras são referidas não tem significado; simbolizam-se por (CNJ).

Se uma qualquer relação de ordem (>) for imposta ao conjunto então o símbolo do conjunto ordenado será (CNJ>):

Se um dado conjunto for um elemento de um produto cartesiano, a ordem stricta imposta nesse conjunto resulta da ordem do produto cartesiano. Esses conjuntos dizem-se multiplicidades e simbolizam-se por (MULT).

O conjunto de tipo (CNJ) de todas as frases simples com o mesmo verbo diz-se uma correspondência ou relação e simboliza-se por (REL).

O conjunto (CNJ) de todas as primeiras palavras (PLV1) de uma relação (REL) diz-se domínio da relação e simboliza-se por (DOM). Identicamente as terceiras palavras (PLV3) numa relação (REL) formam o contradomínio e simboliza-se por (CON).

Os conjuntos (CNJ) podem ser definidos empregando as connectivas neles impostas.

É necessário prover o sistema de instrumentos de processamento tendo em atenção a estrutura algébrica e topológica do domínio e contradomínio do verbo, bem como a natureza deste.

B5) Relações (REL)

A linguagem formal (F_LING) assenta no conceito de frase holonica (HOL) que consiste num triplo estritamente ordenado de palavras colhidas no dicionário (DIC), e.g.: (plv1, plv2, plv3), onde (plv1, plv3) é um par ordenado e (plv3) é o símbolo de um conjunto, eventualmente vago (fuzzy) de pares ordenados, a que o par (plv1, plv3) pertence.

Otipode (plv2) designou-se por 'verbo' (VRB) e (plv1) invocou (plv3) por intermédio de plv2.

Formalmente, (plv2) configura uma relação (ou correspondência) formal, onde (plv1) é o primeiro membro e (plv3) é o segundo membro do par ordenado membro, eventualmente vago (fuzzy), da relação plv2.

Assim é possível definir o conjunto Domínio (DOM<plv2>):

DOM<plv2> =def= {plv, i plv, i é primeiro membro de um par ordenado que pertence ao conjunto de pares representados por plv2},

Identicamente se define Contra_Domínio (CON<plv2>) :

CON<plv2> =def= {plv, k plv, k é segundo membro de um par ordenado que pertence ao conjunto de pares representados por plv2},

Porque (F_LING) é um dialecto evolutivo duma linguagem Natural (N_LING) igualmente evolutiva, os pares que pertencem a plv2 variam com o tempo.

Sendo o afluxo de comunicações finito, a escala humana, o conjunto de pares (plv2) é finito e possui um cardinal. No Cap. C, são dados alguns esclarecimentos complementares sobre verbos.

B6) Resumo

Os três problemas principais a resolver em (F_LING), são:

- * TRADUÇÃO (TRD) da linguagem natural (N_LING) em (F_LING) e respectiva retro versão (Cap. C).
- ** INFERÊNCIA (INF) facultar (F_LING) da propriedade de, deduzir teoremas (Cap. D).
- *** INDUÇÃO (IND) dotar (F_LING) de meios para estender a sua operacionalidade para além do conteúdo informativo resultante da interacção com os dialogantes externos, humanos ou não, (Cap. E).

Nota Final: (F_LING) estando associada a uma linguagem natural e essencialmente viva, tem de ser 'necessariamente' evolutiva para acompanhar a linguagem natural mas também porque (F_LING) vai 'crescendo' e 'aprendendo' a medida que vai interagindo com os dialogantes externos e assim incrementando o seu património informativo e a extensão do respectivo dicionário.
Eventualmente poderá ampliar o seu património de conceitos inatos.

C) Tradução

C1) Introdução

A tradução da linguagem natural (N_LING) em (F_LING) é possível, em princípio, uma vez que o tipo verbo (VRB) pode representar qualquer conjunto de pares ordenados.

Assim, qualquer estrutura algébrica e topológica tem cabimento, e haverá apenas que reproduzir a respectiva estrutura.

São de referir nomeadamente: Lógicas predicativa e relacional, estruturas: grupos, anéis, módulos, vectores, etc., operando sobre conjuntos que são n_multiplicidades, funções, etc.

Repare-se que a base factua de F_LING são triplas: (plv1, plv2, plv3) e que as estruturas referidas podem ser testadas, quanto à sua verossemelhança, com a realidade.

No domínio da 'tradução' o que interessa é se, na passagem de (N_LING) para (F_LING), o sentido da frase natural não foi "traído".

Para descrever o modo como é executado a tradução da linguagem natural (N_LING) para (F_LING), foi escolhido o método da apresentação de um conjunto de casos típicos e descrição dos 'procedimentos' usados.

Porque a linguagem é essencialmente evolutiva, o número, bem como a forma dos 'procedimentos' vai evoluindo por força das novas comunicações do (F_DIALOG) com o mundo exterior.

C2) Alguns Procedimentos Típicos

* Reconhecimento no Dicionário (Procedimento P0):

Reconhecimento das palavras empregadas na 'frase natural (N_FRS) por meio de uma 'busca' no dicionário (DIC).

Caso a palavra não se encontre em (DIC) ou encontrando-se tenha outro sentido então há que a incluir em (DIC), o que envolve:

- escrever a palavra, eventualmente acrescentando algumas indicações suplementares para a distinguir de outras preexistentes.
- indicar o tipo (TIP): verbo, nome, atributo, etc.
- acrescentar comentários semânticos se necessário.

* Credibilidade das Palavras e Frases (Procedimento P1):

Quantificação da Credibilidade da frase proferida pelo autor. O autor é questionado com o objectivo de declarar o grau de credibilidade da frase natural dum modo geral e das palavras constituintes. Em geral, oferece-se o intervalo [0,1], correspondendo: 0 a 'total ignorância' e 1 a 'total certeza'. O próprio autor é julgado por um "painel" (PANL) e é-lhe atribuído um grau de credibilidade que pode ser variável de frase para frase.

Ao termo dos procedimentos P0 e P1 o (N_REC) da frase natural toma o seguinte aspecto:

	h_	a_Autor	f_Frase	1_plv1	2_plv2	3_plv3	4_plv4
palavra	*	joão	f001	luis	bateu	mário	forte
tipo	*	nome	frase	nome	verbo	nome	attribution
valor	.526	.9	.95	.8	.9	.95	.9

* Sequências Sujeito-Predicado (Procedimento P2):

Esclarecimento de sequências tais como "mário forte". A palavra "forte" é apelido, alcunha ou atributo?

Se for atributo então "mário forte" está por "mário é forte" e a N_1-frase "luís bateu mário forte", desdobra-se em duas frases holônicas: "luís bateu mário" e "mário é forte". Se for apelido então "forte" é designação de um conjunto e poderá construir-se a frase holônica:

"mário pertence (família) forte".

Se for alcunha então é razoável que "mário forte" seja substituído pela palavra "mário_forte".

	h	a	f	1	2	3
palavra:	f18_1	ru1	f18	manuel	bater	mário
tipo:	frase	nome	frase	nome	verbo	nome
valor:	0.0	.9	0.0	.8	.9	.8

palavra:	f18_2	ru1	f18	joão	bater	mário
tipo:	frase	nome	frase	nome	verbo	nome
valor:	0.0	.9	0.0	.7	.8	.9

palavra:	f18_3	ru1	f18	f18_1	ymplyca	f18_2
tipo:	frase	nome	frase	holon	verbo	holon
valor:	1.0	.9	.85	0.0	.9	0.0

* Coplativas (Procedimento P3):

Frases do tipo: "joão, luís e manuel bateram mário".

Dois interpretações são possíveis:

1 - a frase significa que "joão bateu mário", luís bateu mário" e "manuel bateu mário", e então a frase inicial pode ser substituída por três frases holônicas equivalentes.

2 - ou a "triada" (joão, luís, manuel) bateu mário, e então trata-se de um n-múltiplo ordenado. A frase inicial será traduzida por: "joão pertence_a mult_jlm", "luís pertence_a mult_jlm", "manuel pertence_a mult_jlm" e "mult_jlm bateu mário".

* Disjuntivas (Procedimento P4)

Frases do tipo: "joão, luís ou manuel bateram mário".

1 - Só um bateu!?, então a frase é equivalente a:
"se joão bateu mário então [luís e manuel não bateram mário],
"se luís bateu mário então [...]",
"se manuel bateu mário então [...]".

2 - A outra interpretação corresponde ao instanciador < pelo menos um >.

3 - Finalmente pode ainda interpretar-se a frase como < pelo menos um, mas não todos >.

* Implicação, se ... então (Procedimento P5):

Este procedimento ocupa-se do silogismo hipotético "se então" como em: "se manuel bateu mário então joão bateu mário". Criando a palavra "ymplyca", a frase pode ser substituída por (manuel bater mário) ymplyca (joão bater mário).

"ymplyca" é o primeiro exemplo de um verbo que tem por Domínio uma frase e Contradomínio igualmente outra frase, assim "ymplyca" configura um operador funcional.

A holonização da frase dada produz 3 frases holônicas cujas (REC) tomam as formas seguintes:

Note-se que o valor da frase f18_2 depende do valor da frase f18_1.

* Formas verbais (procedimento P6):

Das três palavras de um (HOL) apenas o verbo não pode faltar.

O verbo pode configurar as seguintes figuras típicas:

1: representa uma relação tal como este vocabulário é entendido par ordenado pertencente a relação plv2.

2: representa um "estado" para o sujeito. (plv1 plv2 ...), como em: "AA corre", "BB anda".

Cujo sentido pode ser descrito pelas frases: "AA executa o acto de correr" ou "AA esta no estado de corrida". Contudo existem frases como "joão bateu", que podem significar: "joão bateu num_xxx" e "xxx" é um ente desconhecido, então esta frase pertence ao tipo 1

3: representa um "estado" mas o sujeito é omissivo, como em: "chove", que pode entender-se como: "o_estado_do_tempo é chuva". Poderia reconduzir-se o tipo 3 ao tipo2 se fosse explicitado o sujeito.

As relações (formais) estão contidas no conceito de verbo mas há frases onde o verbo generaliza o conceito de relação, muito embora possam ser construídas frases do tipo (HOL) que tenham igual significado mas forma diferente.

* Formas ad_verbais (Procedimento P7):

O verbo, em linguagem natural, muda de forma para atender: ao tempo (passado, presente, futuro).

ao modo (fortemente, habilmente, bem, mal).

a quantidade (muito, pouco,).

ao local (aqui, em Lisboa, em casa,). etc. etc.

O tratamento formal destas formas "modeladores" do verbo pode ser idêntico ao tratamento predicativo das palavras: plv1 e plv2 numa dada frase. Por exemplo:

"passado" representa um intervalo de tempo, aberto à esquerda e à direita. Assim, "João correu", significa que: "correr pertence ao passado" e "passado" é um "predicado" de correr e "predicado" tem o sentido dado em lógica.

"em Lisboa" também é um "predicado" para um verbo, como em: "viver em Lisboa", "correr em Lisboa" etc.

O processo de modelação dos verbos consiste em (F_LING), tratá-los pelo emprego de lógicas predicativas, tal como se procede com plv1 e plv3.

Pode instituir-se uma taxonomia para os "predicados" dos verbos, e.g.: "predicados" de tempo, local, modo, etc. e depois desdobrar cada um destes conceitos.

Estas taxonomias não são exigidas sintacticamente.

Os "predicados" dos verbos podem ser munidos de connectivas, como qualquer conjunto, e assim autorizar operações formais sobre eles.

Assim, todas as palavras que representem substantivos, atributos, verbos ou frases podem ser membros de conjuntos usuais ou vagos (fuzzy).

Todos os conjuntos podem ser munidos de estruturas algébricas e ou topológicas, relações de ordem, medidas etc. desde que, semanticamente, tal seja permitido.

* Interpretação de Frases Naturais (Procedimento P8):

Em geral, as frases naturais são compostas e possuem ou não explicitamente muitos verbos. O procedimento da interpretação representa uma das operações essenciais e que da origem a um diálogo pesado entre (H_DIALOG) e (F_DIALOG). Aquisição referida apenas as operações mais importantes:

- > Análise semântica da frase com a identificação do verbo principal e dos secundários pela ordem da inclusão.
- > Identificação das implicações múltiplas.
- > Introduzir parenteses na frase natural a fim de configurar expressão matemática.

Atingida esta forma, declara-se a frase interpretada.

* Desarticulação ou Holonização da frase (Procedimento P9):

Porque em (F_LING) as frases (HOL) são normalizadas, há que traduzir a frase interpretada (com os parenteses) em tantas frases (HOL) quantas as necessárias. Designando esta operação por "tradução". A holonização pode entender-se como uma tradução em (F_LING) de frases em (N_LING) mas já interpretadas.

Nota final ao Cap. C:

Muitos outros procedimentos seriam de referir mas os apresentados parecem ser suficientes para ilustrar o tema.

A necessidade de tratar frases naturais cada vez mais complexas, que a 'vivência' do (F_DIALOG) com (H_DIALOG) vai impondo, obriga a constantes revisões dos procedimentos já construídos e desenhar novos procedimentos.

Este processo, em princípio, prosseguirá indefinidamente, acompanhado a evolução da linguagem natural e os progressos da própria (F_LING) na sua capacidade de constituir uma boa imagem de (N_LING).

D) Inferência

D1) Introdução

Convém recordar que toda a informação foi tipificada, valorizada e interpretada.

Admite-se que todos os procedimentos que constituem o sistema de Tradução operaram correctamente e que foi obtida uma imagem holonizada das frases naturais razoavelmente correcta.

Em (F_LING), o "filtro principal" foi instalado logo à entrada, de modo a evitar a introdução de inconsistências ou ambiguidades da linguagem natural susceptíveis de ser "sanáveis" pela via do diálogo.

A valorizações são entendidas como 'índices de pertença' de caracterizantes de conjuntos vagos (fuzzy sets).

A construção de conjuntos vagos fez-se usando as connectivas definidas no reticulado de apoio.

nos exemplos representados são: max e min e o conjunto universal do reticulado é: [0, 1].

Porque as frases holónicas (HOL) não são, por definição e construção, não susceptíveis de decomposição em frases ainda mais simples, toda a inferência resultará:

- ou da agregação de frases holónicas;
- ou da desarticulação das frases holónicas em palavras;
- com a formação ou não de conjuntos de palavras.

As necessidades de inferência são assim satisfeitas com três operadores (ou procedimentos):

- Um Meta_Operador de Especificação
- Um Operador_Construtor de Conjuntos
- Um Operador_Construtor de Frases Agregadas

Convém que estes operadores sejam desenhados com a base estrutural de conjuntos vagos (fuzzy sets), uma vez que se dispõe da informação respectiva, isto é, dos 'índices de pertença'.

Recorda-se que os conjuntos Vagos tem como caso particular os conjuntos usuais e assim é sempre possível passar dos primeiros para os segundos.

D2) Meta-Operador de ESPECIFICAÇÃO

Para evitar a reprodução aqui da teoria dos conjuntos vagos, o tema vai ser introduzido com um exemplo simples. Vamos supor que (F_LING) só possui, no presente estado, a informação abaixo, que consta de 5 (HOL):

plv_h	plv_a	plv_f	plv_1	plv_2	plv_3	
f14_1 frase 0.0	joão nome .9	f14 frase 0.0	rui nome .85	bateu verbo .95	mário nome .9	HOL_1
f14_2 frase 0.0	joão nome .9	f14 frase 0.0	rui nome .85	ser verbo 1	forte nome .8	HOL_2
f14_3 frase 0.0	joão nome .9	f14 frase 0.0	f14_1 nome 0.0	ymplyca verbo 1	f14_2 nome 0.0	HOL_3
f12 frase .67	luis nome .95	f12 frase .9	rui nome .88	bateu verbo .99	joão nome .9	HOL_4
f10 frase .576	luis nome .8	f10 frase .9	rui nome .85	bateu verbo .96	mário nome .98	HOL_5

É fácil reconhecer que os 3 primeiros (HOL) resultam da holonização de um silogismo hipotético (se... então) e os restantes dois são simples declarações.

Exemplos de especificações:

Exemplo 1: Especificar o conjunto dos AUTORES (frs_1) que se referem a "rui" com um índice de pertença (I.P.) superior ou igual a .80.
Resposta do meta_operador: plv_1:: (joão, luis)

Exemplo 2: Idem mas em frases com um I.P. > .85. Resposta do meta_operador: plv_1:: (luis).

Exemplo 3: Especificar as frases (HOL) onde ocorre: plv_4 = 'bater' e plv_5 = 'mário'. Resposta do meta_operador: f14_1 e f10. Mas se for imposta uma credibilidade mínima = .5 então a resposta só incluirá: f10.

* Procedimento (MET_SPCFC):
(meta_especificador);

Recorda-se que o (REC) de uma frase holonica (HOL), escrita em (F_LING) tem:
6 palavras (as três últimas formam o (HOL)),
6 tipos (classes de palavras) e 6 valores
(índices de pertença de conj. Vagos)

O especificador terá de ser desenhado em termos de, a 'busca' no banco de frases (HOL) poder ser feito com [1 a 18] especificações singulares arrumadas em 3 classes:

- classe 1 POSição(plv_dado, plv_x) > 0 ou = 0;
- classe 2 POSição(tip_dado, tip_y) > 0 ou = 0;
- classe 3 vlr_dadoA >= vlr[z] e ou
vlr_dadoB <= vlr[w];

onde:

x, y, z, w :: [h.a.f.1..3];

..._dado. :: palavras, tipos ou valores;

Encontradas as frases (HOL) dá-se por terminada a função do meta_especificador e o comando passa ao procedimento seguinte: "Construtor de Conjuntos".

D3) Procedimento (CONS_CONJ):
(construtor de conjuntos);

O operador meta_especificador termina a sua tarefa indicando os (REC) que satisfazem a uma dada especificação. Mas um (REC) holónico tem 6 palavras, 6 tipos, 6 valores, será pois necessário, para construir o conjunto, informar que palavras, tipos e ou valores são de reter, isto é, o conjunto a formar pode ser tão extenso quanto o (REC) ou tão reduzido como o simples numero_ordem do (REC).

Exemplos:

1 Reter uma palavra e o respectivo tipo e valor. O conjunto a construir será essencialmente de triplos (plv, tip, vlr).

2 Reter um par de palavras e os tipos e valores respectivos. O conjunto será de pares de triplos ((plvA, tipA, vlrA), ..., (plvN, tipN, vlrN)).

3 Reter n_uplos de palavras, tipos e valores. ((plvA, tipAvlrA), ..., (plvN, tipN, vlrN)).

4 Reter n_uplos de palavras e valores. Como em: ((plvA, vlrA), (plvB, vlrB), ..., (plvN, vlrN)).

5 Reter 1_uplos de palavras. Como em: (plvA, plvB, ..., plvN).

Porque a caracterização dum conjunto assenta na respectiva especificação, será conveniente conservar esta especificação como parte essencial para a boa compreensão do conjunto. A estrutura do (REC_CONJ) constará da especificação do conjunto, da informação a reter e da lista dos n_uplos retidos.

a) Especificação do conjunto (INST1).

Será agora oportuno referir que a (F_LING) possui uma linguagem auxiliar para receber instruções (F_INSTR). Esta linguagem é um dialecto da linguagem natural (N_LING) e de cada palavra natural conserva os 4 primeiros caracteres. No dicionário (DIC) podem ser incluídas estas abreviaturas. (F_INSTR) é também um dialecto de (F_LING). As instruções (INST1) destinadas ao Operador de Meta.

Especificação tomam a forma de um 'string' que contém uma frase em (N_LING), eventualmente com as palavras truncadas aos 4 primeiros caracteres. Um interpretador converte (INST1) em instruções para o Operador Destinatário.

A especificação do conjunto será justamente a referida instrução (INST1).

b) Informação a reter (INST2).

O objectivo é registar em outro 'string' (INST2) as instruções relativas as informações a reter com o fim de construir o conjunto.

A linguagem será igualmente (F_INSTR).

c) Lista dos n -uplos retidos (UPLOS);
O número de ordem (ORD) dos (REC) escolhidos é sempre memorizada a fim de ser facilitado o acesso à informação residente em disco.
Há dois procedimentos possíveis:
ou a lista se resume aos (ORD), (LIST_ORD).
ou a lista contém as informações instruídas por (INST2) e que será designada por (LIST_COMPL).

Sugere-se que se opte pela (LIST_ORD) e só se passa a (LIST_COMPL) quando necessário

A estrutura do (REC_CONJ) na forma mais usual será:

INST1 : configurado por um string
[160] escrito em F_INST;
INST2 : configurado por um string
[160] escrito em F_INST;
ORD : um array [0..N] de inteiros,
com N:: [10 a 50];

Desejando-se formar uma (LIST_COMPL) haverá que acrescentar:

PLV : um array [0..N, 0..6] of string [201];
TIP : um array [0..N, 0..6] of string [201];
VLR : um array [0..N, 0..6] of real;
onde N :: [10 a 50].

D4) Construtor de Frases Agregadas

O objectivo é construir frases em linguagem natural (N_LING) a partir do conjunto de (HOL) retidos e escritos em (F_LING).

A forma mais simples será 'escrever' sucessivamente as frases pertencentes ao conjunto, deixando a interpretação ao dialogante humano.

Pode contudo munir-se (F_DIALOG) com procedimentos destinados a realizar operações tais como:

- separar o texto da frase holonica (plv1, plv2, plv3) das restantes palavras, tipos e valores.
- referir em seguida aos autores e os símbolos das frases de que são responsáveis.
- tratar o tema das valorizações no fim.

D5) Resumo

O processo de inferência não cria nada de novo que não exista já ou como informação contida no Dicionário (DIC) ou nos recordos (REC) de frases holónicas (HOL).

As saídas resultam das instruções recebidas do (H_DIALOG), isto é, do exterior ao (F_DIALOG).

A resposta do (F_DIALOG) é uma função das instruções (INST) recebidas e da informação contida no dicionário (DIC_STT) e no património informativo (PTR_STT), os quais variam no tempo uma vez que se trata de linguagens evolutivas, donde que as respostas não sejam invariantes ou uniformes.

E) Indução

E1) Introdução;

A indução é a função humana (e dos seres animados em geral) mais difícil de emular.

Alguns exemplos simples descrevem o processo da indução:

a) Suponhamos que, no património informativo (PTR), há várias ocorrências do verbo 'casar_com'.

- Reconhecer que nunca ocorre a forma :
'yyy casou_com xxx' 'casar_com' não é uma relação reflexiva.

- Reconhecer que a forma 'xxx casou_com yyy' não impede que ocorra a forma: 'xxx casou_com xxx', o que seria um sintoma de se tratar de uma relação simétrica.

- Se existir uma aplicação do conjunto S dos seres xxx e yyy num conjunto Z, essa aplicação vai permitir a construção de card (Z) classes de equivalência. Se $\text{card}(Z) = 2$, com sucede com a propriedade sexualidade, haverá duas classes e será possível experimentar o conceito da equivalência e verificar que:
(xxx pertence classe1) :: (yyy pertence classe2)
e concluir que a forma verbal: casado_com, corresponde ao conceito hetero_sexualidade, e que os pares: (xxx,yyy) e (yyy,xxx) cujos elementos são colhidos em classes de equivalência distintos.

b) Em (PTR) há ocorrências dos verbos 'embater' e 'bater'.

- Reconhecer igualmente a simetria e a irreflexibilidade, em 'embater'.

- Mas 'bater' já pode ser reflexivo e, em geral, não é simétrico.

c) Em (PTR) há ocorrências de 'xxx ser_mais_alto_do_que yyy'

- Reconhecer uma relação de ordem estrita na forma verbal:
'ser_mais_alto_do_que'.

d) Em (PTR) há referências a 'correr', 'nadar', 'saltar' e 'chover', etc.

- Reconhecer que os três primeiros verbos são relações com domínio (DOM), mas é omissa o contradomínio. O quarto verbo, 'chover', seria uma relação onde o domínio e o contra_domínio estariam omissos. Como regularizar estes casos?, declarar que nem todos os verbos são relações? ou estender o conceito de relação?.

Os exemplos que seguem procuram ilustrar algumas operações no domínio da indução.

Verificada em a) a simetria de 'casar_com', em 50 casos por exemplo, será legítimo aplicar o conceito de simetria ao (HOL) contido em (PTR), seguinte: "jbc casou_com mft"? Se for legítima essa aplicação, então acrescentar ao (PTR) o (HOL) seguinte: 'mft casou_com jbc'.

Quantos casos são requeridos para justificar o enunciado de uma regra?, 50, 500, 5000? e se houver excepções? estas invalidam a regra ou aguardam a indução de uma regra mais geral que as inclua no seu domínio. Notar que se for exigida um número elevado de casos, então é muito lenta a indução de regras.

Quem 'imagina' as regras a experimentar (simetria, reflexibilidade, linearidade, distância, etc.).

Se forem considerados como conceitos "inatos" nos seres animados reais, nesse caso será necessário habilitar, à partida, o (F_DIALOG) com essas regras.

Em (I_LING) deverão existir os procedimentos semânticos correspondentes.

Existindo nos seres animados reais a faculdade de aprender novas regras de indução, será necessário dotar o (F_DIALOG) com instrumentos de aprendizagem de reconhecimento e aprendizagem de regras de indução, e.g.:

simetria, antisimetria, reflexibilidade, linearidade, verticalidade, horizontalidade, ortogonalidade, distância, relação de ordem, continuidade, triangularidade, circularidade, etc.

Com regularidade ou aleatoriamente ou por instrução externa o património informativo é 'vasculhado' com o objectivo de verificar se certas propriedades podem ser associadas a certos conjuntos de (IOL), e assim induzir novas regras.

E2) Conceito de FORMAS PURAS ou PARADIGMAS;

Suponhamos que (F_LING) possui um procedimento que o habilita a verificar se um dado conjunto foi ou não dotado de uma 'relação de ordem estrita' (r.o.e.). O procedimento poderá ser 'inato' ou 'adquirido'. No (PTR_INF) de (F_LING) existe um verbo (VRBabc) de que há N ocorrências (e.g.: 'ser mais alto que').

Aplicando o procedimento (r.o.e.) as N ocorrências verificam-se poucas excepções. Nesta circunstância será justificado criar, (imaginar) um verbo paradigmático (VRBabc_P) que só contém os (HOL) de (VRBabc) que satisfazem a (r.o.e.) e qualquer futuro (IOL) de (VRBabc) que 'passe o filtro' (r.o.e.) pertencerá a (VRBabc_P).

A vantagem da criação do paradigma (VRBabc_P) está justamente no facto deste possuir uma "propriedade" (a relação de ordem estrita, r.o.e.) que só parcialmente se verifica em (VRBabc). São essas propriedades que definem o paradigma. Assim foram imaginadas: as formas geométricas puras, as estruturas algébricas e topológicas, as lógicas, etc. Todas satisfazem, com rigor, a propriedades bem definidas.

E3) Proximidade

Sendo os 'paradigmas' formas ideais abstraídas da realidade, passou a ser imperativo que se encontrasse um processo para apreciar o "afastamento" entre um dado 'real' e um 'paradigma' de modo a habilitar o (F_DIALOG) a aceitar o 'paradigma' como uma razoável 'aproximação' do 'real' e daí atribuir a este as propriedades do paradigma, e.g.:

Julgada uma dada figura suficientemente 'próxima' de um triângulo, declarar que a área dessa figura é igual à área do triângulo (ideal) que a emula.

Do exposto resulta que a cada 'paradigma' corresponderá um procedimento que avalia a 'PROximidade' entre um dado 'real' e o paradigma, sem este procedimento de pouco teria servido a criação do paradigma.

A colecção de procedimentos 'PROX' associada à colecção de paradigmas, constitui a faculdade básica do 'reconhecimento'. Mesmo que (F_DIALOG) não seja criativo terá pelo menos ser dotado da capacidade de 'reconhecer'.

E4) Comentários Finais

A forma como foi introduzido o conceito de paradigma, permite ainda proceder da seguinte forma (forma a-priori): Definir um verbo_paradigmático (relação), possuindo um conjunto de propriedades entre si compatíveis.

Este verbo_paradigma poderá constituir uma criação ocasional ou inata ou vinda do exterior e pode entender-se como um 'paradigma a priori', uma vez que foi proposto antes de terem sido 'observados', no património (PTR_LING), casos que justifiquem ou permitam a proposição ou criação desse verbo_paradigmático.

A forma (a-posterior) resultará da indução propriamente dita.

Parece assim razoável que (F_DIALOG) seja dotado de:

- 1 Uma colecção de paradigmas e respectivos procedimentos de verificação de propriedades e medidas de proximidades. Constituiria um património racional innato (PRI_DIALOG).
- 2 Faculdade de aprendizagem no sentido de ser possível acrescentar o (M_DIALOG) que desempenharia a função de Tutor.
- 3 Faculdade de, espontaneamente, acrescentar ou alterar (PRI_DIALOG).

Das três faculdades, é a última que está hoje na fronteira da investigação e para a qual não há ainda uma resposta satisfatória.

Espera-se que a holonização do diálogo possa vir a permitir atacar com sucesso esta terceira faculdade.

O património racional inato corresponderá biologicamente ao conjunto cromosómico que todo ser vivo recebe quando "nasce" e que o habilita passar em poucos anos ou dias, de uma simples célula a um ser adulto, com todos os atributos nomeadamente: aprendizagem e gerar novos seres semelhantes.

Alguns exercícios foram já programados em Pascal, para demonstrar o conceito, como por exemplo:

Relação de Ordem Estrita: que verifica a existência de uma (r.o.e.) na forma verbal "ser_mais_alto_do_que".

Distância em grafos de relações. Proximidades (entre imagens paradigmáticas e "reais").

PALAVRAS RESERVADAS

DIALOG

transferência de comunicações entre: Humanos, Instrumentos e Natureza em geral. Um linguagem é instrumental nesse DIALOG.

II_DIALO

Dialogante Humano.

F_DIALOG

Um instrumento formal de diálogo.

PRI_DIALOG

Património Racional inato de F_LING.

LING

Linguagem.

N_LING

Linguagem Natural (língua viva em geral).

F_LING

Linguagem Artificial e em geral formal usada internamente por F_DIALOG.

F_INST

Dialecto abreviado de N_LING e F_LING, usando palavras truncadas (Primeiras 4 letras) e destinado a instruções e comandos.

INSTx

Instrução escrita em F_INST.

DIALECT

Qualquer linguagem LING associada ou derivada duma linguagem natural N_LING.

F_DIALECT

Corresponde ao F_LING.

DIC

Dicionário Universal, onde estão contidas quer as palavras da N_LING como as de F_LING.

DIC_STT

Porque as linguagens são evolutivas, convém dar uma designação ao 'Estado' em que se encontra o dicionário DIC: DIC_STT.

FRSzzz

O radical FRS significa que se trata de uma 'frase'.

N_FRS

Trata-se de uma frase em N_LING (linguagem natural).

F_FRS ou FRS

a frase está em F_LING.

HOL

Porque as frases em F_LING possuem uma forma normalizada que se designa de 'holonica', as F_FRS ou FRS também se designam por: HOL.

IIOL*

Frase holonica inversa de HOL.

PLVxx

O radical PLV significa que se trata de uma palavra contida no dicionário DIC ou numa frase.

TIPxx

O radical TIP significa o 'tipo' a que pertence uma dada palavra PLV. Supõe-se a existência de uma 'estrutura taxonomica de tipos'.

VL.Rxx

O radical VL.R representa o 'índice de pertença' que mede a 'credibilidade' que merece a palavra PLV. No caso da F_LING apresentada, VLR toma valores no intervalo [0, 1] da recta real, onde 0 significa

AUTyy

NOME do 'Autor' da frase ou declaração.

VRBxx

Trata-se de uma classe de tipos designada usualmente por 'verbos'. Nas frases holónicas HOL ocupa a posição central da triada de palavras.

REC

Significa recorde (record), trata-se de uma forma normalizada de registar frases.

PTR_INF

Representa o 'património informativo' da linguagem F_LING num dado estado, ou seja, o conjunto de todos os HOL retidos.

PTR_STT

Simboliza o estado de património PTR_INF.

CNJ

Símbolo de conjunto, quer seja no sentido usual quer no sentido conjunto_vago (fuzzy).

I.P.

Índice de pertença (caracterizante dum Conjunto Vago).

CNJ

Conjunto ordenado.

MULT

Elemento de um produto cartesiano. As palavras num MULT estão ordenadas pela ordem do produto cartesiano.

REL

Símbolo de relação ou correspondência ou função. Um HOL configura, em geral, uma relação, mas pode ter um sentido mais geral. As relações estão contidas em HOL mas não se verifica o inverso.

LIST_ORD

Conjunto contendo os números de ordem dos recordes dum conjunto de frases holónicas de um certo tipo, por exemplo: conjunto de (HOL) com o mesmo verbo.

LIST_COMPL

Idêntico a LIST_ORD, mas contendo a transcrição parcial ou total do conteúdo do (HOL).

DOM

Domínio da relação REL.

CON

Contra_domínio (range) da relação REL.

TRD

Significa 'tradução'. São operações de versão e retro_versão entre as linguagens N_LING e F_LING.

INF

Significa 'inferência'. Este vocábulo é reservado para descrever operações o património PTR_INF, que não envolvem, para o F_DIALOG, 'criar' ou 'imaginar' hipóteses ou produzir 'informação' que não exista já no PTR_INF.

MET_SPCFC

Operador ou procedimento destinado a especificar conjuntos, cumprindo instruções recebidas em F_INST (dialecto de F_LING).

ORD

Num. de ordem dum recorde (record) numa fila (file).

LIST_ORD

Listagem dos números de ordem dos (REC) apurados.

MET_SPCFC

Meta teorema da especificação.

PAANL

Painel de "humanos" que apreciam os autores.