Gul

{LING2.000}
{ling2.sum}

SUMMARY

F_LING is a formal language based on the concept of holonic phrase (or HOL), a 4 word sentence, where the first word symbolises the semantic environment .
The holonic phrase , preceded by 2 words (the symbols of the initial phrase and of the HOL), is memorised in a repository REP , the constancy of the number of words (2+4) in each sentence is very helpfull.
The main purpose is to show how induction can be accomplished.

SOMMAIRE

F_LING est un langage formel base' sur le concept de phrase holonic (ou HOL), une phrase qui a 4 mots , dont le premier symbolise le contexte semantic .
La phrase holonic , precedee de 2 mots (les symbols de la phrase initial et du HOL), et memorisee dans une base de phrases REP , la constance du nombre de mots (2+4) de chaque phrase est tres util .
L'objectif principal est de montrer l'apptitude de F_LING a' l'induction .

SUMARIO

F_LING e' uma linguagem formal baseada no conceito de frase holonica (ou HOL) , uma frase com 4 palavras , sendo a primeira o simbolo do contexto semantico .
A frase holonica, precedida de 2 palavras (os simbolos da frase inicial e do HOL), e' memorisada num repositorio REP , a constancia do numero de palavras (2+4) de cada frase e' muito util .
O principal objectivo e' mostrar a aptidao de F_LING a' induccao .

* F_LING *

Autor : Antonio G.Portela (I.S.T.) .

A) Introduction.

1 * F_LING is a formal language based on the concept of "holonic phrase" , (HOL) , a 4 word declaration (tetrad),

which is considered the simplest structure conveying some information e.g. :

HOL_z PlvO Plv1 Plv2 Plv3

The "holonic phrase" , HOL_z , informs :

- there is a "nexus" between the words Plv1 and Plv3 , and (Plv1 , Plv3) is an oriented pair of words .
- the "nexus" is named Plv2 and , in general ,there are many other pairs having the same nexus .
- the triad Plv1 Plv2 Plv3 is valid in the semantic context or referencial , symbolised by PlvO .

Example: HOL_z Jack Mary is_maried_to John .

- " HOL_z " is the symbol of the "holonic phrase" .
- " Jack " has made the following declaration ,
  " Mary is maried to John " .
- " Jack " is also the symbol given to the semantic context .
- there is a "nexus" between " Mary " and " John " .
- (" Mary " , " John ") is an oriented pair .
- " is_maried_to " is the word that symbolises the above refered nexus .

Note that the characters S D V C are used sometimes in the text instead of the indices O, 1, 2, 3 .

2 * F_DIALOG (viz. Ref.5) is a liguistic instrument that enables the dialog with F_LING , using idiomatic language instead of a specific one .

The main functions of F_DIALOG are :

- Elimination of ambiguities .
  A dialog with the user takes place and terminates when F_DIALOG considers the "ambiguous idiomatic phrase" clarified .
  The dialog may not take place if the idiomatic phrase is "clear" , not ambiguous , from the point of view of F_DIALOG.
- De-articulation of the "clarified phrase" in one or more "holonic phrases".
  This operation is performed without loss of "information" and it is possible to revert (reconstruct) an idiomatic phrase with the same "information content" of the original clarified idiomatic phrase .
- Placing all the "holonic phrases" in a "repository" (REP) . The standart form of an HOL in REP is :
  Frs_x Hol_z PlvO Plv1 Plv2 Plv3
  - Frs_x is the symbol of the "clarified idiomatic phrase".
  - Hol_z is the symbol of one of the "holonic phrases" derived from Frs_x by F_DIALOG.
- Creating a diccionary (DIC) , with all the words given to or created by F_DIALOG .
- Revise REP and DIC whenever necessary .
  Both , REP and DIC are "indefinitely" expansible .
- Finaly , F_DIALOG "asks" how sure PLVO is about the statement made and , if possible , a value in an appropriate "credible_scale" is requested .
  F_DIALOG "asks" also a "credible_value" regarding PlvO (the semantic context) .

3 * Examples :

Example 1 :

*Frs_A :: "the coordinates of point P are 235 , 29, 4 "
After clarification of Frs_A , one obtains :
Frs_a :: "the Base of the space is Base_0.
"235, 29 and 4 correspond , in Base_0 , to the second , third and first coordinates , respectively ."

After de-articulation of Frs_a , the following "holonic phrases " are obtained .

Frs_a	Hol_1	Base_0	4	is	first_coord.
Frs_a	Hol_2	Base_0	235	is	second_coord.
Frsa_a	Hol_3	Base_0	29	is	third_coord.

*Frs_B :: " Base_0 is a base for a 3 dimensional linear ortogonal space (3_Orto_Space) and all coordinates are given in meters ".

The de_articulation , of Frs_B gives :

Frs_B	Hol_4	3_Orto_Space	Base_0	is	a_Base
Frs_B	Hol_5	Base_0	meter	is	1_coord_unit
Frs_B	Hol_6	Base_0	meter	is	2_coord_unit
Frs_B	Hol_7	Base_0	meter	is	3_coord_unit

Example 2 :

Frs_J1 "John said that Peter is tall"
Frs_A4 "Anton said that Peter is small"
Frs_P1 "Anthropologist said that John is a pygmy"
Frs_Q3 " ??? Anton plays basketball "

Can be converted as follows :

Frs_J1	John_point_of_view	Peter	is	tall
Frs_A4	Anton_point_of_view	Peter	is	small
Frs_P1	Antrop._statement	John	is	a_pygmy
Frs_Q3	X ??	Anton	plays	basketball

Example 3 :

Frs_X2 " Mary said if John is gambling then his spending is great "

Converts to :

Frs_X2	Hol_24	Mary_statem.	John	is	gambling
		80	0	0	0
Frs_X2	Hol_25	Mary_statem.	John_spend.	is	great
		80	0	0	0
Frs_X2	Hol_26	Mary_statem.	Hol_24	implies	Hol_25
		80	0	75	0

Obs.: "Mary" is highly credible (80) .
The implication (Hol_26) is judged by "Mary" as highly credible (75) .
But the credible_value of both Hol_24 and Hol_25 is not given by Mary in Frs_x2 .

B) REPOSITORIES .

All information, converted in holonic phrases (HOL) , is archived in the repository (REP) and all words employed are listed in the dictionary (DIC) , both can be indefinitely increased and revised .
Several types of sub-repositories can be extracted from REP, selecting in REP all the holonic phrases, HOLs, that possesse a given set of "words" in a given set of "places" in the HOL .
The normalised form of an holonic frase in REP is :

FRS_ HOL_ PLV.S_ PLV.D_ PLV.V_ PLV.C_

where :

FRS_ reference to the original phrase .
HOL_ simbol of the holonic phrase .
PLV.S_ semantic context.

PLV.D_ element of the domain of the "nexus" .
PLV.V_ designation of the "nexus" .
PLV.C_ element of the range of the "nexus" .

1) Type REP_X (1 constrain) .
   There are 4 distinct cases and depending on de meaning of X , (X : S, D, V, C) , the following sets are given :
   - REP_S : the semantic contexts PLV.S_ .
   - REP_D : the domain elements PLV.D_ .
   - REP_V : the verbs PLV.V_ .
   - REP_C : the range elements PLV.C_ .
2) Type REP_X_Y (2 constrains) .
   There are 6 distinct cases , e.g. :
   REP_S_D , REP_S_V , REP_S_C , REP_D_V , REP_D_C , REP_V_C.
   Note that : REP_X_Y = REP_Y_X and REP_X_X is meaningless.
3) Type REP_X_Y_Z (3 constrains).
   There are 4 distinct cases , e.g. :
   REP_S_D_V , REP_S_D_C , REP_S_V_C , REP_D_V_C .
4) Type REP_X_Y_Z_W (4 constrains) .
   There is at most 1 case that corresponds to a well defined HOL but this type of sub_repository may be an empty set .

Some sub_rep are much used and should be "constructed" before starting solving a given problem , as for example :

REP_S (with : S = S1...Sn) :
This type is useful to solve problems , where only the set of semantic contexts S1...Sn are of interest . It may be more expedite to create, previously, a sub_rep of the type REP_S .

REP_S_V (with : V = Vv and S = Ss , St) :
This type of sub_rep. may be useful to study the verb Vv in the semantic contexts Ss and St .

REP_S_V_C (with : C = Cc , V = Vv and S = Sa...Sq):
To endow the set Dv (domain of Vv) with a topologic structure , this type of sub_repositories are well addapted .

All these sub_rep. , may configurate a "list of addreses" of the holonic phrases (HOL) archived in REP .
The amount of memory used is much reduced but the access to REP is inderect .

C) Primitive and Derived concepts .

The primitive concepts in F_LING are the holonic-phrases HOL and the words PLV , which are archived respectively in the repository REP and the dictionary DIC .
Concepts like set, multiples, elements of a set etc. are derived concepts and a presentation of some typical derived concepts is appropriate :

1) SET (CJT) .
All sub_repositories presented in Cap B are sub_sets of REP and REP is the "universal_set" of HOL .
The void (or empty) set and REP are sets .
The diccionary DIC is the "universal_set" of words (PLV) and there is a word "is_member_of" , a verb, with which a very important type of HOL can be created , namely :

FRS_f HOL_h D_x is_a_member_of C_c .

Building the REP based on D_x is_member_of C_c , all D_x retained configurate a set, eventualy an empty one, that is a sub_set of the diccionary DIC .
The symbol of a set is CJT and the above refered set of words can be represented as follows :
CJT(C_c) = {D_x : D_n ,....., D_w}

or = {D_x : FRS_f HOL_h D_x is_a_member_of C_c} , the first mode is a simple listing of words , the second is based on the operator "is_a_member_of" that enables the construction of the set .

2) RELATIONS (REL) .
A relation can be exhaustively represented by the set of all pairs (nexus between 2 words) that are members of a certain verb Vv , (sub_rep. of the type REP_Vv) .
The list may include nexus in various semantic-environments.
The problem arises when the relation REL has infinite nexus and the exhaustive representation is not feasable .
Infinite sets will be delt in cap. E .
The strict_order_relation (>=) will be used as an example :
Let CJT = { a, b, ... ,n } be a set and >= a verb .
In REP there are the following holonic_phrases :

H_ab b >= a H_ac c >= a .... H_an n >= a
H_bc c >= b H_bd d >= b H_bn n >= b
....
H_mn n >= m

The method expounded avoids the use of "rules" and is based on simple declarations .

3) MULTIPLES (MLT) .
Let CJT>=C be an ordered set of sets CJT_x :
CJT>=C = [{ CJT_a ... CJT_n } >=] ;
A MLT_x of CJT>=C is any ordered set of elements of CJT_x and x : [{ a ...n} >=] .
A MULT is thus an element of the cartesean product of CJT_x of CJT>=C .

4) CONNECTIVES (CNT) .
Connective is the general designation of a class of relations used to endow a set with an algebric structure .
In general a cartesean product is formed between the set and another set (eventualy with the same set) , and a relation is defined , the cartesean product beeing the domain and the set the range , e.g. : nax,min sup,inf +,* union,conjuction etc. etc.
A very simple example, introducing the connective Max, follows : The given set is {0 , 1} .

H_1	S_boole	0,0	is_an_ordered_pair_of	Cart.Prod.
H_2	"	0,1	"	"
H_3	"	1,0	"	"
H_4	"	1,1	"	"
H_5	"	0,0	Max	0
H_6	"	0,1	Max	1
H_7	"	1,0	Max	1
H_8	"	1,1	Max	1

5) FUNCTIONAL Relations (FUNC) .
The domain and or the range represente Relations , as for instance :

Functional	:	D=relation	V=Funct.	C=real
Transform	:	D=relation	V=Transf.	C=relation
Distribution	:	D=set	V=Distr.	C=relation

6) OPERATORS (OPER) .
The standard form of a relation REL is a "table" of all the nexus which are members of REL (D_dom. V_verb C_range).
Another form of describing REL is by means of operatores :
direct : (OP_REL) , OP_REL { D_ } >>> C_

inverse : (OP_REL!) , OP_REL! { C_ } >>> D_
If these operators are given then the above refered "table" , is dispensable .
The inverse relations are usefull in the topological structering of the domain set of REL .

D) TRANSLATION (TRAD) .
The basic concept of HOL is described by an ordered set of 4 words, a tetrad, (S_ D_ V_ C_) , the first one S_ symbolises the semantic environment or referential, the remaining triad (D_ V_ C_) is a declaration .
If the referential or semantic environment D_ is substituted by d_ , then the triad (D_ V_ C_) is converted or "translated" to (d_ v_ c_) , with the objective of conveying the same information :

(S_ D_ V_ C_) == (s_ d_ v_ c_)

A point P in a plane (R2) , is spatialy described by two reals (coordinates of de point P) .
Let the coordinates be : 20 (coord.1), 300 (coord.2) .
This information would be incomplete , without an explicit mention to the "base" or referencial used , e.g. :

H_1	S_baseA	20	is_coord.1_of	P
H_2	S_baseA	300	is_coord.2_of	P

Two types of translation are here presented :

* Literal or word_to_word translation .

H_11	S_baseA	translates_to	S_baseB
H_12	20	"	40
H_13	300	"	150

H_1 and H_2 are "translated" as follows :

H_1_t	S_baseB	40	is_coord.1_of	P
H_2_t	S_baseB	150	is_coord.2_of	P

* Semantic or phrase_to_phrase translation .

H_s1	H_1	translates_to	H_1_t
H_s2	H_2	translates_to	H_2_t

In general , these two methods yield diferent results , but are equivalent, in formal languages .
The inverse translation does not reproduce the inicial holonic phrases , in general .
Again it is considered importante the existence of an inverse operation in formal languages .

E) INFINITE SETS (continuum) .
Finite processors are unable to deal with the continuum and the usual method consists in the conversion of the continuum in a discrete form .
A table of circular functions is a discrete image of the continuous functions it represents .
The output of a program that computes a circular function is also discrete .
The choice between using more memory or consuming more time in computation is an economical problem but not a conceptual one .
The "coarsenness" of a table can be reduced by interpolators.
In the following , the method of obtaining a discrete image of the continuum will be considered adequate to the type of problem to be solved and this assumption is applicable both to tables and operators (hard or soft) .

F) CREDIBILITY .
Credibility is fundamental in F_LING .
As this mater is expounded in ref. 1 , a sumarised reference is deemed sufficient .

To the tetrad of words corresponds a tetrad of reals belonging to the closed set [-1 , 1] .
The absolute value measures the credibility :
0 (crebility nil) ; 1 (Total credibility)
Negative numbers stand for "negation" , e.g. :
- 1 means totaly false .
A set of functions over the set of tetrads yields the credibility of the set .

G) DEDUCTION & INDUCTION .

G1) Introductory example :

*Deduction .

The verb "is_married_to" (i_m_t) is declared a symmetrical relation e.g.: if (X i_m_t Y) then (Y i_m_t X) or H_1 X i_m_t Y ; H_2 Y i_m_t X ; H_12 H_1 implies H_2 . Thus, the declaration : (A i_m_t B) and the symmetry of the (i_m_t) relation , permits to infer : (B i_m_t A) . This is a typical deduction , the total information in REP has not been increased , a theorem demonstration was performed .

*Induction .

A process of induction applied to the study of properties of the verb (i_m_t) may be described by the following succession of steps :

1 Data Retrieving regarding (i_m_t) .
Find in REP, all the HOLs where (i_m_t) is the verb .
The following list ,(A) , was obtained :
A1 i_m_t B6 , A2 i_m_t B5 , A4 i_m_t B3 ,
B5 i_m_t A2 , B7 i_m_t A3 , B3 i_m_t A4 .

2 Looking for Symmetries in list A .
The processor being endowed with the capacity and the initiative to test various concepts, namely the symmetry of relations , yields the list (B) of all pairs of symmetric HOLs contained in list (A) :
((A2 i_m_t B5) and (B5 i_m_t A2)) ,
((A4 i_m_t B3) and (B3 i_m_t A4)) .

3 Conjecture propounding .
The 2 cases of symmetry observed and the non existence of counter-examples , may be considered enough to posit the following conjecture : (i_m_t) is symmetric .

4 Testing the Conjecture .
From now on, every new HOL containing (i_m_t) as a verb will be tested for symmetry (step 2) .
If some new cases are detected and no counter_examples have arisen , then the conjecture is converted in a "finding" and (i_m_t) is declared a symmetrical verb and a new rule is included in REP .

*After step 4 , the symmetry of (i_m_t) , now a rule , can be used in deductive processes whenever needed .
Eventualy, new findings may imply the revision of the rule.

The development of the theme , inductive processes ,follows .

G2 Endowment of faculties .

All living things (biots) inherit and or aquire "faculties" essencial to perform the species tasks .
The construction of a computer language addapted to a certain species must take in account these faculties .
It would be senseless to provide the computer language with means to describe 3-dimensional objects , if the biots of the species are blind and deaf .
But if a given species is endowed with the faculty of sensing the magnetic inclination , the language should

provide concepts like : "up and down" , distance and velocity and the necessary rules to process the information aquired and appropriate to an one dimentional linear space, with the usual topology and the usual operators : order_relation , connectives (+ , *), scalar derivatives , etc..

Humans are endowed with so many faculties by heritage and further developed by learning, that evan a modern natural language and all formal language so far developed are insufficient to emulate a human being, adequately .

The method usualy adopted is to provide a basic set of formal procedures and functions , e.g. :

- A set of characteres with a finite cardinal and a strict order relation and 2 connectives .
- A set of numbers with a finite cardinal and two not closed connectives (+ , *) , a mimic of the integer numbers .
- A set of numbers , emulating imperfectly the reals .
- A set of predefined functions , on the reals .
- Input and output devises .

The basic set of procedures and rules can be increased, including other specialised functions and or structured sets .

This final set (basic and specialised) represents the "inherited faculties" of the artifact (processor) and emulates a certain "species" real or imagined .

G3 Paradigms and Reconnoitre .

Formal algebic and topological structures and concepts were developed mostly because they serve as models or paradigms .

Natural and formal languages are essencially an antropomorphic artifact , adjusted to and inspired in the faculties and atributes of the human species .

Paradigms can be confronted or compared with the atributes of the real or imaginary objects and beings . e.g. :
counting == cardinal of set ; natural numbers .
precedence (in a tribe or table) == order_relation .
stereocopic vision == distance, eccart, proximity .
cup to contain a liquid == concave (convex).
straight line, vertical, horizontal == linear spaces .
biologic clocks, sideral time == time.

The atributes of a paradigm are invariant and all theorems and inferences are also invariably valid .

If an "object" conforms well with a paradigm then "all" atributes of the paradigm can be applied to the "object" .

The pair (paradigm , to_"recnnoitre"_it) represents a faculty , inherited or learned , by the artifact and the two are usefull only as a pair .

Step 2 , looking for symmetries , is only feasable if the processor is endowed with the pair :
(paradigm symmetry , procedure to reconnoitre symmetry)

G4 Conjectures propounding (step_3) .

The artifact (processor) , at each stage of evolution , possesses a set of pairs (Sprp) , of the form :
(Paradigm , Reconnoitre procedure) or (P,RP)

In REP , there are holonic phrases (HOL) that can be compared with a Paradigm (P_k) by means of a Reconnoitre procedure (RP_k) , the result is given by an element of the range of (RP_k) .
Usualy the range(RP_k) is the set (Yes , No) , but other sets can be contemplated , e.g. : (yes, undecided, no) .

Systhematicaly or randomly , the processor takes the iniciative to proceed with a reconnaissance of a pair (P_k,RP_k), belonging to the set (Sprp) .

The successes and unsuccesses are registered and a decision functor (FC) is provided , to enable the processor to propound the paradigm P_k as a conjecture .

G5 Conjectures testing (step 4) .

Step 4 is mainly designed to test the conjecture , giving time to confirm or infirm the conjecture .

A functor (FT) is provided to perform this function .

Step 4 can be eliminated if the functor (FC) , refered in G4 , and functor (FT) are composed in one (FCT) . (FCT) declares the finding a "rule" , jumping step 3 .

G7 Another example .

This example presents another aspect of the induction processe and is based on simple generalisation or extention of the domain of application of a rule .

John sees a dog(B) bite Jack and declares :

H_23 John dog(B) bites Jack .

The functor (FCT) may suggest "rules" , like :

dog(B) bites Humans , or
all dogs bite Jack , or
all dogs bite all humans , etc. etc.

The generalisation is a "risk avoider" posture , that has saved many human lives .

The functor declares universal what as been observed once .

Eventualy, later, the rule may be disproved so many times, that dogs have to be classified in bad and good dogs and the domain of the rule is reduced to bad dogs .

G7 Conclusion .

The final result of steps 1 to 4 is the proclamation of a "new rule" by the artifact (processor) . This rule was not given by an external source , namely by the human operator .

A regular application of the inductive method , using all the pairs (P_k, RP_k) members of Srpr , will generate eventualy more "rules" , following the introduction of new holonic phrases (HOLs) in the repository REP .

If the repository is "frosen" , that is no "new HOLs" are intruced in REP , then , after some time , the production of "new rules" stops .

Human progress is a long story of disproved conjectures :

The Earth is flat and is in the center of the world.
The space dimention is : 3, later 4, now more then 4 .
The atom , an undividable body, is dividable .
etc. etc.

The statement "all rules are true" is acceptable , provided, the domain of the rule may be an empty set ! .

H) Conclusion .

Humans do not create "ab nihil" , the same applies to human artifacts and to F_LING in particular .

* F_LING starts with a Repository REP_O , a Diccionary DIC_O , a set of hard_ware capacities HW_O .

In REP_O are archived the initial holonic phrases (HOLs) :

. HOLs describing Sprp_O , a set of pairs (Paradigm , Reconnoiter_procedures) .
. Hols describing rules (functions , procedures etc.).

DIC_O contains some initial words .

* The development of F_LING results of the contact with the "outside world" , by means of the input_output devices .
  The information received , after interpretation , is stored in REP .
  By induction, new rules are created , increasing the avaiable information .
  If needed , declarations not contained explicitely in REP can be deducted from the avaiable data , using the rules thar are in force .
  F_LING can also be developed by introducing in Sprp_0 new pairs (P_, RP_) , that will increase the induction capacity of the processor and new rules will be uncovered .

I am gratefull to Professors Martins, Morgado and H.Coelho for their comments but the responsibility for all errors and misconceptions ,the paper may contain , is mine .

REFERENCES .

1 N.BOURBAKI . Actualites Scientifiques et Industrielles , 1957.

2 B.RUSSEL, Introduction to Mathematical Philosophy , George Allen & Unwin Ltd, 1930 .

3 L.A.ZADEH, Fuzzy Sets , Information and Control , 8, 338 - 353 , 1965 .

4 M.EYTAN, Fuzzy Sets : A Topos_Logical Point of View, Fuzzy sets and Systems, Vol. 5, n.1, 47-67 , 1981 .

5 A.G.PORTELA , Dialogos , Inteligencia Artificial , 1988, vol.0, No.2 .