ChatScript System Variables and Engine-defined Concepts

- Engine-defined Concepts
- System Variables
- Control over Input
- Interchange Variables

Engine-defined concepts

In addition to concepts defined in script files, the system automatically defines a bunch of dictionary-based sets as well as dynamically computed concept members.

set	description
~web_url	word is a web url
~email_url	word is an email address
~kindergarten	word learned early in life
~grade1_2	word learned in these grades
~grade3_4	word learned in these grades
~grade_5-6	word learned in these grades.
	Unmarked words are learned
	even later
~utf8	word has nonascii characters
~daynumber	word could be a number of a day
	in a month
~yearnumber	word could be the number of a
	recent year
~dateinfo	phrase is month day year of
	some kind
~kelvin	temperature marker
~celcius	temperature marker
~fahrenheit	temperature marker
~twitter_name	twitter user name
~hashtag_label	twitter topic reference

Interjections, "discourse acts", and concept sets

Some words and phrases have interpretations based on whether they are at sentence start or not. E.g., good day, mate and It is a good day are different for good day.

Likewise sure and I am sure are different.

Words that have a different meaning at the start of a sentence are commonly called interjections.

In ChatScript these are defined by the livedata/interjections.txt file. In addition, the file augments this concept with "discourse acts", phrases that are like an interjection. All interjections and discourse acts map to concept sets, which come thru as the user input instead of what they wrote.

For example yes and sure and of course are all treated as meaning the discourse act of agreement in the interjections file. So you don't see yes, I will go coming out of the engine.

The interjections file will remap that to the sentence \sim yes, breaking off that into its own sentence, followed by I will qo as a new sentence.

These generic interjections (which are open to author control via interjections.txt) are:

interjection	description
~yes	
~no	
~emomaybe	
~emohello	
~emogoodbye	
~emohowzit	
~emothanks	
~emolaugh	
~emohappy	
~emosad	
~emosurprise	
$ ilde{\ }$ emomisunderstand	
~emoskeptic	
~emoignorance	
~emobeg	
~emobored	
~emopain	
~emoangry	
~emocurse	
~emodisgust	
~emoprotest	

interjection	description
~emoapology ~emomutual	

Because all interjections at the start of a sentence are broken off into their own sentence, this kind of pattern does not work:

u: (~yes _*)

You cannot capture the rest of the sentence here, because it will be part of the next sentence instead. This means interjections act somewhat differently from other concepts.

If you use a word in a pattern which may get remapped on input, the script compiler will issue a warning. Likely you should use the remapped name instead.

The following concepts are triggered by exactly repeating either the chatbot or oneself (to a repeat count of how often repeated). Repeats are within a recency window of about 20 volleys.

concept	description
~repeatme	
~repeatinput1	
~repeatinput2	
~repeatinput3	
~repeatinput4	
~repeatinput5	
~repeatinput6	

POS (Part of Speech) Tags

Words will have pos-tags attached, specififying both generic and specific tag attributes, eg., ~noun, ~noun_singular.

Generic Specifics

nouns	description
~noun	
~noun_singular	
~noun_plural	
~noun_proper_singular	
~noun_proper_plural	

nouns	description
~noun_gerund	
~noun_number	
~noun_infinitive	
~noun_omitted_adjective	
verbs	description
~verb	
~verb_present	
~verb_present_3ps	
~verb_infinitive	
~verb_present_participle	
~verb_past	
~verb_past_participle	
~aux_verb	
~aux_verb_present	
~aux_verb_past	
~aux_verb_future	
~aux_verb_tenses	
~aux_be	
~aux_have	
~aux_do	

Auxilliary verbs are segmented into normal ones and special ones. Normal ones give their tense directly. Special ones give their root word. The tense of the be/have/do verbs can be had via `properties() and testing for verb tenses

adjectives	description
~adjective	
~adjective_normal	
~adjective_number	
~adjective_noun	
~adjective_participle	
adjectives in comparative form	n description
~more_form~most_form	
~adverb	
~adverb_normal	

adverbs in comparative form	description
~more_form~most_form	
~pronoun~pronoun_subject~pronoun_object	
~conjunction_bits~conjunction_coordinate~conjunction_subordinate	
~determiner_bits~determiner~pronoun_possessive~predeterminer	
~possessive	covers 'and 's at end of word
~to_infinitive	"to" when used before a noun
~preposition~particle	free-floating preposition tied to
~comma	
~quote	covers 'and _"_ when not en
~paren	covers opening and closing par
~foreign_word	some unknown word
~there_existential	the word there used existentia

In addition to normal generic kinds of pos tags, words which are serving a pos-tag role different from their putative word type are marked as members of the major tag they act as part of. E.g,

	description
~noun_gerund	verb used as a
	~noun
~noun_infinitive	verb used as a
	~noun
~noun_omitted_adjective	an adjective
	used as a
	collective noun
	(eg the beautiful
	$are \ kind)$
~adjectival_noun	noun used as
	adjective like
	bank "bank
	teller"
~adjective_participle	verb participle
	used as an
	adjective

For <code>~noun_gerund</code> in *I like swimming* the verb gerund *swimming* is treated as a noun (hence called noun-gerund) but retains verb sense when matching keywords tagged with part-of-speech (i.e., it would match <code>swim~v</code> as well as <code>swim~n</code>).

Additionally, there is

	description
~number	is not a part of speech, but is comprise of "noun_number" (a normal number value like 17 or seventeen)
~adjective_number	also a normal numeral value and also ~placenumber) like first.
~integer	
~float	
~positiveinteger	
~negativeinteger	
~modelnumber	not a true number, but a word with both alpha and numeric

To can be a preposition or it can be special. When used in the infinitive phrase To go, it is marked $\sim to_infinitive$ and is followed by $\sim noun_infinitive$.

	description
	description
~verb_infinitive	refers to a
	match on the
	infinitive form
	of the verb $(I$
	hear John sing
	or I will $sing$).
~There_existential	refers to the use
	of where not
	involving
	location,
	meaning the
	existence of, as
	in There is no
	future.
	racaro.

	description
~Particle	refers to a
	preposition
	piece of a
	compound verb
	idiom which
	allows being
	separated from
	the verb. If you
	say I will call
	off the meeting,
	call off is the
	composite verb
	and is a single
	token. But if
	you split it as
	in I will call the
	meeting off,
	then there are
	two tokens.
	The original
	form of the verb
	will be call and
	the canonical
	form of the verb
	will be call_off,
	while the
	free-standing off
	will be labeled
	~particle.
~verb_present	will be used for
	normal present
	verbs not in
	third person
	singular like I
	walk and
~verb_present_3ps	will be used for
	things like he
	walks

	description
~possesive	refers to 's and
	'that indicate
	possession,
	while possessive
	pronouns get
	their own
	labeling
	~pronoun_possessive
~pronoun_subject	is a pronoun
	used as a
	subject (like he)
~pronoun_object	refers to
	objective form
	like him

Individual words serve roles in the parse of a sentence, which are retrievable. These include:

	description
~mainsubject	
~mainverb	
~mainindirect	
~maindirect	
~subject2	
~verb2	
~indirectobject2	
~object2	
~subject_complement	adjective object
	of sentence
	involving
	linking verb
~object_complement	2ndary noun or
	infinitive verb
	filling
	modifying
	mainobject or
	object2
	<pre>act_verb~conjunct_adjective~conjunct_adverb~conjunct_phr</pre>
~postnominalAdjective	·
	occuring
	AFTER the
	noun it
	modified

	description
~reflexive	reflexive
	pronouns
~not	
~address	noun used as
	addressee of
	sentence
~appositive	noun restating
	and modifying
	prior noun
~absolutephrase	special phrase
	describing
	whole sentence
~omittedtimeprep	modified time
	word used as
	phrase but
	lacking
	preposition
	$(Next\ tuesday\ I$
	$will\ go)$
~phrase	a prepositional
	phrase start
	(except
~clause	a subordinate
	clause start
~verbal	a verb phrase

and special concepts: $| \sim capacronym |$ word is in all caps (and &) and is likely an acronym

System Variables

The system has some predefined variables which you can generally test and use but not normally assign to. These all begin with %. Ones that are reasonable to set are written in bold underline. Boolean values are always 1 or null on returns. 1 or 0 if you are setting them.

Date & Time & Numbers

variable	description
%date	one or two digit day of the month

variable	description
%day	Sunday, etc
%daynumber	r0-6 where $0 = Sunday$
%fulltime	seconds representing the current time and date
	(Unix epoch time)
%hour	0-23
%timenumbe	erempletely consistent full time info in numbers
	that you can do _0 =
	^burst(%timenumbers) to get _0 = seconds
	(2digit) _1=minutes (2digit) _2=hours (2digit)
	_3=dayinweek(0-6 Sunday=0) _4=dateinmonth
(1-31) _5=month(0-11 January=0) _6=year.Ye	
	need to get it simultaneously if you want to do
	accurate things with current time, since
	retrieving %hour %minute separately allows
	time to change between calls
%leapyear	boolean if current year is a leap year
%daylights	savoings if current within daylight savings
%minute	0-59
%month	1-12 (January = 1)
%monthname	January, etc
%second	0-59
%volleyting	neumber of seconds of computation since volley
	input started
%time	hh:mm in military 24-hour time
%week	1-5 (week of the month)
%year	e.g., 2011
%rand	get a random number from 1 to 100 inclusive

Time and date information are normally local, relative to the system clock of the machine CS is running on. See \$cs_utcoffset for adjusting time based on relationship to utc (e.g your server is in Virginia and you are in Colorado).

User Input

variable	description
%bot	current
	bot
	responding

variable	description
%revisedinput	Boolean
	is
	$\operatorname{current}$
	input
	$_{ m from}$
	input
	not
	direct
	$_{ m from}$
	user
%command	Boolean
	was the
	user
	input a
	command
%foreign	Boolean
	is bulk
	of the
	sen-
	tence
	com-
	posed
	of
	foreign
	words
%impliedyou	Boolean
- •	was the
	user
	input
	having
	you as
	implied
	$\operatorname{subject}$
%input	the
-	count
	of the
	number
	of
	volleys
	this
	user
	has
	made
	ever
	· · · -

variable	description
%ip	ip
	address
	supplied
%language	current
	dictio-
	nary
	language
%length	the
	length
	in
	tokens
	of the
	current
	sentence
%more	Boolean
	is there
	another
	sen-
	tence
	after
	this
%morequestion	Boolean
	is there
	a? or
	ques-
	tion
	word in
	the
	pend-
	ing
	sentences

variable	description
%originalinput	all sen-
	tences
	user
	passed
	into
	volley,
	before
	ad-
	justed
	in any
	way
	except
	OOB
	data is
	stripped
	off
%originalsentence	the
J	current
	sen-
	tence
	after to-
	keniza-
	tion but
	before
	any
	adjustments
%parsed	Boolean
•	was
	current
	input
	parsed
	successfully
%question	Boolean
•	was the
	user
	input a
	ques-
	tion -
	same as
	? in a
	pattern
	•

variable	description
%quotation	Boolean
	is
	current
	input a
	quotation
%sentence	Boolean
	does it
	seem
	like a
	sen-
	tence
	(sub-
	ject/verb
	or
	command)
%tense	past,
	present,
	or
	future
	$_{\rm simple}$
	tense
	(present
	perfect
	is a
	past
	tense)
%user	user
	\log in
	name
	$\operatorname{supplied}$
${\tt \%userfirstline}$	value of
	%input
	that is
	at the
	start of
	this
	conver-
	sation
	start

variable	description
%userinput	Boolean
_	is the
	current
	input
	from
	the user
	(vs the
	chatbot)
%voice	active
	or
	passive
	on
	current
	input

Chatbot Output

variable	description
%inputrejoinderag of	
	any pending
	rejoinder for
	input or null
	if none
	pending
%lastout	puthe text of
	the last
	generated
	response for
	the current
	volley -
	always null
	across volleys
%lastque	st B molean did
	last output
	end in a?
%outputr	ej ouih edteargif
	system set a
	rejoinder for
	its current
	output or 0

variable	description
%response	number of
	committed
	responses
	that have
	been
	generated for
	this sentence
	(see
	Advanced
	User-
	Advanced
	Output:
	Committed
	Responses

System variables

variable	description
%all	Boolean
	is the
	:all flag
	on?
	(:all to
	set)
%document	Boolean
	is :docu-
	ment
	running
%fact	Numeric
	value
	most
	recent
	fact id
%freetext	kb of
	avail-
	able
	text
	space

```
variable
             description
%freedict number
             of
             unused
             dictio-
             nary
             words
% free fact number
             of
             unused
             facts
%maxmatchvanigiladsites
             \operatorname{number}
             of
             match
             vari-
             ables,
             cur-
             rently
             20
\verb|\maxfactset| \textbf{bis} ghest
             number
             of
             @fact-
             sets,
             cur-
             rently
             20
%host
             name of
             the
             current
             host
             ma-
             chine or
             "local"
{\tt \%regression} Boolean
             is the
             regres-
             sion
             flag on
```

variable	description
%server	Boolean
	is the
	system
	running
	in
	server
	mode
%rule	get a
	tag to
	the
	current
	execut-
	ing rule.
	Can be
	used in
	place of
	a label

variable	description
%topic	name of
	the
	current
	"real"
	topic .
	if
	control
	is cur-
	rently
	in a
	topic or
	called
	from a
	topic
	which is
	not
	system
	or
	nostay,
	then
	that is
	the
	topic.
	Other-
	wise the
	most
	recent
	pend-
	ing
	topic is
	found
%actualtor	
	the
	current
	topic
	being
	pro-
	cessed
	(system
	or not)

variable	description
%trace	Numeric
,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	value of
	the
	trace
	flag
	(:trace
	to set)
%httpresp	,
	code of
	most
	recent
	^jsonopen
	call
%pid	Linux
	process
	id or 0
	for
	other
	systems
%restart	You
	can set
	and
	retrieve
	a value
	here
	across a
	system
	restart.
%timeout	Boolean
	tells if a
	timeout
	has
	hap-
	pened,
	based
	on the
	time-
	limit
	com-
	mand
	line
	parameter

Build data

variable	description
%dict %engine %os	date/time the dictionary was built date/time the engine was compiled os invovled (linux windows mac ios)
%script %version	date/time build1 was compiled engine version number

You actually can assign to any of them. This will override them and make them return what you tell them to and is a particularly BAD thing to do if this is running on a server since it affects all users (unless you reset the variable at the end of the volley. Assigning a period to a variable resets it).

Typically one does this as a temporary assignment in a #! comment line to set up conditions for testing using :verify.

Making them return a new value is NOT the same thing as making the engine have a different value. Unless the variable is marked as settable, setting a value affects only the value returned by a future call to the system variable. It does not change engine values the variable is meant to reflect.

Control Over Input

The system can do a number of standard processing on user input, including spell correction, proper-name merging, expanding contractions etc. This is managed by setting the user variable \$cs_token.

The default one that comes with Harry is:

The #signals a named constant from the dictionarySystem.h file. One can set the following:

These enable various LIVEDATA files to perform substitutions on input:

flag	description	
#DO_ESSENTIALS	perform LIVEDATA/systemessentials which	
	mostly strips off trailing punctuation and sets	
	corresponding flags instead	
#DO_SUBSTITUTES	perform LIVEDATA/substitutes	
#DO_CONTRACTIONS	perform LIVEDATA/contractions, expanding	
	contractions	
#DO_INTERJECTIONS	perform LIVEDATA/interjections, changing	
	phrases to interjections	
#DO_BRITISH	perform LIVEDATA/british, respelling brit words	
	to American	
#DO_SPELLING	performs the LIVEDATA/spelling file (manual	
	spell correction)	
#DO_TEXTING	performs the LIVEDATA/texting file (expand	
	texting notation)	
#DO_SUBSTITUTE_SYSTEM	M do all LIVEDATA file expansions	
#DO_INTERJECTION_SPLITTIMEAk off leading interjections into own sentence		
#\$DO_NUMBER_MERGE	merge multiple word numbers into one (four and	
	twenty)	
#\$DO_PROPERNAME_MERGE	merge multiple proper name into one (George	
	Harrison)	
#DO_DATE_MERGE	merge month day and/or year sequences (January	
	2, 1993)	
#JSON_DIRECT_FROM_OOB	asking the tokenizer to directly process OOB data.	
	See ^jsonparse in JSON manual.	

The contents of the files are pairs of tokens per line. Left is the word to replace and right is the replacement. When multiple words are involved, the left side uses underscores to represent this and the right side uses +. If the right side is missing, it means just delete.

If any of the above items affect the input, they will be echoed as values into <code>%tokenFlags</code> so you can detect they happened. The next changes do not echo into <code>%tokenFlags</code> and relate to grammar of input:

flag	description	
DO_POSTAG	allow pos-tagging (labels like ~noun ~verb become marked)	
DO_PARSE	allow parser (labels for word roles like ~main_subject)	
DO_CONDITIONAL_POSTACperform pos-tagging only if all words are known.		
	Avoids wasting time on foreign sentences in particular	
${\tt NO_CONDITIONAL_IDIOM\ will\ not\ perform\ substitutions\ in\ the\ dictionary\ which}$		
	are considered conditional idioms	

flag	description
NO_ERASE	where a substitution would delete a word entirely as junk, don't
DO_SPLIT_UNDERSCORES	happens after all other input tokenization and adjustments except number merge, and separates words that have been conjoined either because the dictionary has them (<code>credit_card</code>) or because they were merged by proper name merging, or by substitution. The result is only words without underscores (excluding number words like <code>five_thousand_and_four</code>
MARK_LOWER	if a word is considered a proper name in CS and is marked as an upper case word, this will force it to perform any markings for its lower case form as well. Sometimes users type stuff in upper case that really should be lower

Normally the system tries to outguess the user, who cannot be trusted to use correct punctuation or casing or spelling. These block that:

flag	description
STRICT	_EXASEING
	for 1st
	word of
	a sen-
	tence,
	assume
	user
	uses
	correct
	casing
	on
	words

```
flag
         description
NO_INFERDQUESTION
         system
         \ will\ not
         try to
         set the
         QUES-
         TION-
         MARK
         flag if
         the user
         didn't
         input a
         ? and
         the
         struc-
         ture of
         the
         input
         looks
         like a
         question
DO_SPELDEMECKO
         internal
         \operatorname{spell}
         checking
ONLY_LOWERCASE
         input
         (except
         "I") to
         be
         lower
         case,
         refuse
         to rec-
         ognize
         upper-
         case
         forms
         of
\begin{array}{c} \text{anything} \\ \text{NO\_IMPERATIVE} \end{array}
NO_WITHIN
```

flag description

NO_SENT#NOEDEEND

break
input
into
sentences

Normally the tokenizer breaks apart some kinds of sentences into two. These prevent that:

flag ${\it description}$ ${\tt NO_COLOM}{\underline{\circ}}{\tt END}$ break apart a sentence after a colon ${\tt NO_SEMICOLON_END}$ breakapart a sentence after a semicolonUNTOUCHED: etneut this alone, will tokenize only on spaces, leaving everything but spacing untouched

```
{\tt LEAVE\_QifOTip} ut
        is found
        withing
        " " it
        will
        become
        {\bf a} \ {\bf single}
        token
        exactly\\
        as it is
        seen.
        W/o
        Leave\_Quote,
        it is
        con-
        verted
        into a
        word
        without
        quotes
        and
        using
        under-
        scores
        instead
        of
        spaces.
        So "My
        Fair
        Lady"
        be-
        comes
        My_Fair_Lady,
        which
        would
        match a
        movie
        title if
        you had
        one,
        unlike
        My Fair
        Lady
        becom-
        ing the
        result-
       26ng
        token
        and
```

 ${\it unrecognized}$

description

flag

flag description

Note

you can change \$cs_token on the fly and force input to be reanalyzed via `retry(SENTENCE). I do this when I detect the user is trying to give his name, and many foreign names might be spell-corrected into something wrong and the user is unlikely to misspell his own name.

Just remember to reset \$cs_token back to normal after you are done. Here is one such way, assuming \$stdtoken is set to your normal tokenflags in your bot definition outputmacro:

If you type my name is Rogr into a topic with this, the original input is spell-corrected to my name is Roger, but this will change the \$cs_token over to one without spell correction and redo the sentence, which will now come back with my name is Rogr and be echoed correctly, and \$cs_token reset.

That's assuming nothing else would run differently and trap the response elsewhere. If you were worried about that, it would be possible for the script to save where it is using <code>fgetrule(tag)</code> and modify your control script to return immediate control to here after input processing if you had changed <code>\$cs_token</code>.

Private Substitutions

While in general, substitutions are defined in the LIVEDATA folder, you can define private substitutions for your specific bot using the scripting language. You can say

```
replace: xxx yyyyy
```

which defines a substitution just like a livedata substitution file. It actually creates a substitution file called privateO.txt or private1.txt in your TOPIC folder.

Even then, those substitutions will not be enacted unless you explicitly add to the \$cs_token value #DO_PRIVATE, eg

```
$cs_token = #DO_INTERJECTION_SPLITTING |
    #DO_SUBSTITUTE_SYSTEM |
    #DO_NUMBER_MERGE |
    #DO_PROPERNAME_MERGE |
    #DO_SPELLCHECK |
    #DO_PARSE |
    #DO_PRIVATE
```

The left side of the substitution pair is case insensitive (matches either case on input) and can be placed in double-quotes (which converts spaces to underscores internally).

The right side of the substitution pair is case sensitive and can be placed in double-quotes (which converts spaces to plus signs internally).

Similarly while canonical values of words can be defined in LIVEDATA/SYSTEM/canonical.txt, you can define private canonical values for your bots by using the scripting language. You can say:

```
canon: oh 0 faster fast
```

which defines new canonical values for things and creates a file canon0.txt or canon1.txt in your TOPIC folder.

If you want to set a canonical pair from a table during compilation, you can use a function to do the same thing (but only 1 pair at a time).

```
^canon(word canonicalform)
```

Numeric Substitutions

A special kind of private substitution (equally applicable in regular substitution files) is the numeric substitution.

```
replace: ?_km kilometers
```

The ?_ matches a digit number followed immediately by km, like 1.2km and will separate the number and replace the units with the given replacement. The input can be singular or have an 's' like 10.5dollars. And it can be with or without abbreviation periods, like 10kps or 10k.p.s

Interchange Variables

The following variables can be defined in a script and the engine will react to their contents.

interchange variable	description
\$cs_token	described
	exten-
	sively
	above

```
interchange variable
                      {\it description}
                      controls
$cs_response
                      auto-
                      matic
                      han-
                      dling of
                      outputs
                      to user.
                      By
                      default
                      it
                      consists
                      of
                      $cs_response
                      #Response_upperstart
                      #response_removespacebeforecomma
                      #response_alterunderscores
                      #response_removetilde
                      If you
                      want
                      none of
                      theses,
                      use
                      cs_response
                      =0 (all
                      flags
                      turned
                      off).
                      See
                      ^print
                      for
                      expla-
                      nation
                      of flags.
                      #response_noconvertspecial
                      - leave
                      escaped
                      n r and
                      t alone
                      in
                      output
                      and
                      \log
             30
                      #response_upperstart
                      - makes
                      the first
                      letter of
                      an
                      output
                      sen-
                      tence
```

interchange variable	description
\$cs_jsontimeout	seconds before JsonOpen de- clares a time out failure. If unspeci- fied the default
\$cs_crashmsg	is 300 in server mode, what to say if the server crashes and we return a message to the user. By default the message is Hey, sorry. I forgot what I was thinking
\$cs_abstract	about. used with :abstract

interchange variable	description
\$cs_looplimit	loop()
	defaults
	to 1000
	itera-
	tions
	before
	stop-
	ping.
	You can
	change
	this
	default
	with
	this

interchange var	riable description
\$cs_trace	if this
	variable
	is
	defined,
	then
	when-
	ever the
	user's
	volley is
	fin-
	ished,
	the
	value of
	this
	variable
	is set to
	that of
	:trace
	and :trace is
	cleared
	to 0,
	but
	when
	the user
	is read
	back in,
	the
	:trace is
	set to
	this
	value.
	For a
	server,
	this
	means
	you can
	perform
	tracing
	on a
	user
	w/o
	making all user
	${ m transac} ext{-} { m tions}$
ç	dump
,	trace
	urace

 ${\rm data}$

interchange variable	description
\$cs_control_pre	name of
	topic to
	run in
	gambit
	mode
	on pre-
	pass,
	set by
	author.
	Runs
	before
	any sen-
	tences
	of the
	input
	volley
	are ana-
	lyzed.
	Good
	for
	setting
	up
	\dot{n}
	values
<pre>\$cs_usermessagelin</pre>	nitmax
	number
	of mes-
	sage
	pairs
	(user
	input &
	bot
	output)
	saved
	in topic
	file

interchange variable	descriptio
\$cs_externaltag	name of
	a topic
	to use
	to
	replace
	existing
	internal
	English
	pos-
	parser.
	See
	bottom
	of
	ChatScript
	PosParser
	manual
	for
	details

interchange variable	description
\$cs_prepass	name of
	a topic
	to run
	in re-
	sponder
	mode
	on
	main
	volleys,
	which
	runs
	before
	\$cs_control_main
	and
	after all
	of the
	above
	and
	pos-
	parsing
	is done.
	Used to
	amend
	prepa-
	ration
	data
	coming
	from
	the
	engine.
	You can
	use it
	to add
	your
	own
	spin on
	input
	process-
	ing
	before
	going
	to your
	main
	control.
	I use it
	to, for
36	exam-
	ple,
	label
	com-
	mands
	as ques-
	tions,
	aton

stan-

interchange variable	description
\$cs_control_main	name of
	topic to
	run in
	respon-
	der
	mode
	on
	main
	volleys,
	set by
	author
<pre>\$cs_control_post</pre>	name of
	topic to
	run in
	gambit
	mode
	on post-
	pass,
	set by
ф1	author
\$botprompt	message for
	console
	window
	to label
	bot
	output
\$userprompt	message
, F F -	for
	console
	window
	to label
	user
	input
	line
<pre>\$cs_crashmsg</pre>	message
	to use if
	a server
	crash
	occurs

interchange variable	description
\$cs_language	if
	spanish,
	will
	adjust
	spell
	check-
	ing for
	spanish
	colloquial
\$cs_token	bits
	control-
	ling
	how the
	tok-
	enizer
	works.
	By
	default
	when
	null,
	you get
	all bits
	as-
	sumed
	on. The
	possible
	values
	are in
	src/dictionarySystem.h
	(hunt
	for
	\$token)
	and you
	put a #
	in front
	of them
	to gen-
	erate
	that
	named
	nu-
	meric
	constant

interchange variable	description
\$cs_abstract	topic
	used by
	:ab-
	stract
	to
	display
	facts if
	you
	want
	$_{ m them}$
	displayed
<pre>\$cs_prepass</pre>	topic
	used be-
	tween
	parsing
	and
	$\operatorname{running}$
	user
	control
	script.
	Useful
	to sup-
	plement
	parsing,
	setting
	the
	ques-
	tion
	value,
	and
	revising
	input
	idioms

interchange variable description $cs_{\without model} \$ matchvariable covers

multiple words,

what

should sepa-

rate

them-

by default

it's a

space,

but

under-

score is handy

too.

Initial

system

charac-

ter is space,

creat-

ing

 ${\it fidelity}$

with

what was

typed.

Useful

if $_$ can

be recognized

in input

(web addresses).

Chang-

ing to _

is consistent

with multi-

word

repre-

senta-

tion and

key-

word recogni-

40

interchange variable	description
\$cs_userfactlimit	how
	many of
	the
	most
	recent
	perma-
	nent
	facts
	created
	by the
	script
	in re-
	sponse
	to user
	inputs
	are kept
	for each
	user.
	Std
	default
	is 100
<pre>\$cs_outputchoice</pre>	for
	regres-
	sion:
	forces
	specific
	one of a
	output
	choice
	block -
	base 0
<pre>\$cs_response</pre>	controls
	some
	charac-
	teristics
	of how
	re-
	sponses
	are
	formatted

interchange variable	descripti
\$cs_randIndex	the
	random
	seed for
	this
	vollev

interchange variable	description
cs_utcoffset	if
-	defined,
	then
	%time
	returns
	current
	utc
	time +
	$_{ m time}$ -
	zone
	offset.
	The
	offset is
	usually
	a
	$_{\mathrm{simple}}$
	number,
	mean-
	ing
	hours,
	and can
	have +
	$\operatorname{or}-\operatorname{in}$
	front of
	it. It
	can also
	be a
	normal
	$_{ m refer}$
	ence like
	02:30
	which
	means
	plus 2
	hours
	and 30
	minutes
	beyond
	utc, or -
	01:30:20
	which
	means 1
	hour,
	30 min-
43	utes,
	and 20
	seconds
	before
	utc (as
	if
	anyone

anyone would

interchange variable	description
\$\$db_error	error
	mes-
	sage
	from a
	post-
	gres
	failure
	\$\$find-
	$text_start$
	- ^find-
	text
	return
	the end
	nor-
	mally,
	this is
	where it
	puts
	the
	start
\$\$tcpopen_error	error
	mes-
	sage
	from a
	tcpopen
	error
\$\$document	name of
	the doc-
	ument
	being
	read in
	docu-
	ment
	mode
<pre>\$cs_randindex</pre>	current
	value of
	the
	random
	genera-
	tor
	value

interchange variable	description
\$cs_bot	name of the bot cur- rently in use
\$cs_login	login name of the user
\$\$csmatch_start	start of found words from ^match
\$\$csmatch_end	end of found words from ^match
<pre>\$cs_fullfloat</pre>	if defined, causes the system to gen- erate full float 64-bit preci- sion on out- puts, other- wise you get 2 digit preci- sion by default

interchange variable	description
\$cs_botid	when
_	non-
	zero
	creates
	facts
	and
	func-
	tions
	re-
	stricted
	by this
	bit-
	$\max k$ so
	facts
	and
	func-
	tions
	created
	by
	other
	$_{ m masks}$
	cannot
	be seen.
	allows
	you to
	sepa-
	rate
	facts
	and
	func-
	tions
	per bot
	in a
	multi-
	bot
	environ-
	ment.
	During
	compi-
	lation if
	this is
	set by a
	bot:
	com-
	mand,
	then
46	func-
	tions
	created
	and
	facts
	created

by tables

if
defined,
causes
the
system
to
output
num-
bers in
a differ-
ent
lan-
guage
style:
french,
indian.
All
other
values
are
english.
t if
defined
changes
how
many
times
you can
pass
back
RETRY_TOPIC
before
it fails
(current
limit is
30)
t <u>s</u> ek deded
topic
retry
limit is
encountered

interchange variable	description
\$cs_topicretrylimi	t if
	defined
	changes
	how
	many
	times
	you can
	pass
	back
	RETRY_TOPIC
	before
	it fails
	(current
	limit is
	30)
<pre>\$cs_saveusedJson</pre>	if not
	null,
	the only
	JSON
	facts
	CS will
	write
	into the
	user's
	topic
	files
	that are
	referred
	to (di-
	rectly
	or indi-
	rectly)
	from
	user
	vari- ables
	being
	saved.
	(see
	below)
	perow)

interchange variable	description		
\$cs_proxycredentia	\$cs_proxycredential\$ee		
	^JSONOPEN		
	in		
	JSON		
	manual		
<pre>\$cs_proxyserver</pre>	See		
	^JSONOPEN		
	in		
	JSON		
	manual		
<pre>\$cs_proxymethod</pre>	See		
	^JSONOPEN		
	in		
	$_{ m JSON}$		
	manual		

\$cs_saveusedJson exists as a kind of garbage collection. Nowadays most facts will come from JSON data either from a website or created in script. But keeping on top of deleting obsolete JSON may be overlooked. When this variable is non-null, ChatScript will automatically destroy any JSON fact that cannot trace a JSON fact path back to some user variable. Variables that have as values the name of a JSON object or array automatically protect all JSON facts underneath. JSON references merely within some text string will not protect anything, nor will references from some other non-JSON fact.