

Assessing the Utility of ResearchCyc in Recognizing Textual Entailment¹

Christopher Cox

Department of Computer Science

Stanford University

October 10, 2005

Introduction

This paper analyzes ResearchCyc's utility in a real Natural Language (NL) application and provides the potential ResearchCyc user with a sense of what's in the Knowledge Base and how it might be used. Specifically, we have investigated the feasibility of using ResearchCyc as part of Stanford's system initially built for the PASCAL Recognizing Textual Entailment (RTE) challenge (see links below). The crucial question is the recall of ResearchCyc: how often is there sufficient taxonomic and reasoning information in ResearchCyc for it to be able to complete domain-independent natural language inference tasks.

We have chosen not to use ResearchCyc's parser or NL tools as we both were unable to successfully use CycNL components in the early releases of ResearchCyc, and were more interested in interfacing our NL tools' output with ResearchCyc Knowledge. Our plan is to parse sentences and identify grammatical relations using Stanford's tools, and garner information for inferring textual entailment using ResearchCyc's lexicon, argument-frame mapping, and concept hierarchy which can be plugged into various components of our system.

For readability, ResearchCyc's general "hash-dollar" relations (and those only) are

¹ This report is based on research sponsored by DARPA under agreement number FA8750-040200284. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the author and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of DARPA or the U.S. Government.

italicized, so # $\$genls$ gets written as *genls*. All remarks and comments about what's there or not there are current as of ResearchCyc v1.0.

Relevant Links

The PASCAL Recognizing Textual Entailment Challenge

<http://www.pascal-network.org/Challenges/RTE/>

The Stanford system description: *Robust Textual Inference Using Diverse Knowledge Sources*

<http://nlp.stanford.edu/~manning/papers/rte.pdf>

Our parses and dependency analyses of the RTE dataset:

<http://www.stanford.edu/~rajatr/rte/>

ResearchCyc

<http://researchcyc.cyc.com/>

A Word about ResearchCyc Predicates

We rely on the ResearchCyc lexicon to interpret some or all of our dependency parses as conceptual information. There are several ResearchCyc predicates that make statements about relationships between concepts and words given sense, part of speech and/or subcategorization frame. The most basic of these is *denotation*, and the somewhat similar *semTrans* predicates generally contain more complicated conceptual representations of single words or groups of words. For example, *denotation* assertions look like:

```
((#$denotation #Bat-TheWord #SimpleNoun 0 #Bat-Mammal)
```

```
(#$denotation #Bat-TheWord #SimpleNoun 1 #BaseballBat)
```

```
(#$denotation #Bat-TheWord #Verb 0 #BaseballBatting)
```

while a *nounSemTrans* assertion looks like:

```
(#$nounSemTrans #Bachelor-TheWord 0
```

(#\$and (\$\$isa :NOUN \$\$AdultMalePerson) (\$\$maritalStatus :NOUN \$\$Single))

and a *verbSemTrans* assertion (where *word=Feed-TheWord* *sense=0*,
subcatFrame=DitransitiveNPCompFrame):

(#\$verbSemTrans \$\$Feed-TheWord 0 \$\$DitransitiveNPCompFrame
(\$\$and (\$\$isa :ACTION \$\$FeedingEvent)
(\$\$fromPossessor :ACTION :SUBJECT)
(\$\$objectOfPossessionTransfer :ACTION :OBJECT)
(\$\$toPossessor :ACTION :INDIRECT-OBJECT)))

Denotations are generally useful in cases where proper translations of a word haven't been entered into ResearchCyc and something more cursory is acceptable (see ex. 6).

Crucially there are varying levels of granularity, relational, and definitional information in the lexicon, and each word can be “handled” by any number of such predicates.

The best documentation on ResearchCyc NL tools and lexical mapping is at <http://www.cyc.com/cycdoc/ref/nl.html>.

Once we are in the space of ResearchCyc concepts, a huge number of predicates ostensibly relate these concepts to one another. The *isa* (is a) and *genls* (generalization) predicates, which express hyper/hyponymy relations, are among the most reliably present and robust (the two most common, in fact, with 234,116 and 53,956 assertions, respectively). We have designed a search to explore the space of these relations to determine a path between two concepts in ResearchCyc as a measure of similarity/entailment in the common case where no direct connection exists. Also, the reader will see that the term “spec” gets used to talk about a specialization of a collection (the inverse of *genls*), so *KidnappingSomeone* is a spec of *ActsCommonlyConsideredCriminal*.

20 Examples from the 2005 PASCAL RTE dataset

All the following selections are “true” entailments from the Recognizing Textual Entailment dataset that our current system judged as “false”, but which involve a relatively

small piece of common sense knowledge not available in our baseline system, the knowledge of which is mainly what can be derived from WordNet. We've tried to choose entailment examples that are non-trivial, in that they involve some common sense or lexical "missing link" that our current strategy doesn't detect; and not *too* hard, in that they are straightforward and presumably require relatively few such links. They're listed by type and index, along with the non-trivial and missing link on our side of the inference, which is sometimes trimmed down for clarity.

We present ResearchCyc's capabilities and/or preventative deficiencies with respect to each entailment, and provide generalizations where possible. We also present brief discussions of the nature of general Cyclish relationships as they arise; this should be helpful and understandable for a potential user who hasn't mastered all terminology. Indeed, it may even serve as a useful introduction.

1. QA 591

T: A jury is slated to decide for the first time whether Jack Kevorkian, famed as "Dr. Death," has violated Michigan's assisted-suicide ban, while the state continues to grapple with the issue of what to allow when the ill want to end their pain by ending their lives.

H: Jack Kevorkian is the real name of "Dr. Death".

famed as Dr. Death -> is the real name of Dr. Death.

There is no entry for *Fame-TheWord* in ResearchCyc, so we are immediately prevented from working anywhere with this one. There is an entry for *Famous-TheWord*, which, as an adjective, denotes the concept *Famous*, but this knowledge wouldn't help us to understand the verb "fame".

2. QA 565

T: Soprano's Square: Milan, Italy, home of the famed La Scala opera house, honored soprano Maria Callas on Wednesday when it renamed a new square after the diva.

H: La Scala opera house is located in Milan, Italy.

Milan, home of the La Scala -> La Scala is located in Milan.

The Stanford Parser deals with the appositive well; the dependency output indicates that home is an appositive to Italy, that “La Scala Opera House” is an argument to the PP headed by “of”, which in turn is an argument to the NP headed by home.

To recognize this entailment, we would want a module that translates the meaning of “home of” by identifying that the “home” noun phrase needs to be unpacked. We need to tell ResearchCyc that this instance of “home” corresponds to the *GenitiveFrame*, (i.e part of a Genitive Phrase: nouns in association with a preceding possessive or a following ‘of –PP’) which in this case tells us that the *nounSemTrans* of *Home-TheWord* entails the *residesInDwelling* relation between two arguments to “home”, or that POSSESSOR “La Scala” resides in “Home” (the head NOUN). The appositive dependency informs us that “Home” in this case refers to “Milan”.

In the hypothesis, we have that the *verbSemTrans* of *Locate-TheWord* entails the *objectFoundInLocation* between the verb’s SUBJECT and OBJECT slots.

Given this coarse translation of both sentences, we look for relationships between the predicate-argument statements entailed by our translation. In this case, ResearchCyc does not have an obvious relationship between our crucially informative predicates: *objectFoundInLocation* and *residesInDwelling*.

We could conceivably leverage the *genTemplate* pred. in ResearchCyc, which generates more common and perhaps more statistically relevant English paraphrases given a template: in this case, telling us that “permanently located in” is the best paraphrase of *usualLocationOfObject*, the *genIPred* of *residesInDwelling*, and that “located in” is in fact the best paraphrase of *objectFoundInLocation*. We could imagine the general situation where words that aren’t telling us much (“home” in this case) get semantic translations in ResearchCyc, from which we search the space of *genls* (generalizations) towards *semTrans*’ (semantic translations) of target head verbs or predicates with matching arguments. In the case where no ResearchCyc relationship is clear (as above), we use *genTemplate* to translate back into English as we traverse and plug paraphrases back into Stanford’s system at a cost as common paraphrase substitutions of less-informative predicates. Then again, there isn’t an extensive paraphrase database and it’s not clear how robust this method would be, but the sparsity would at least trim down the search space.

3. QA 594

T: For the first time in history, the players are investing their own money to ensure the future of the game," Atlanta Brave pitcher Tom Glavine said.

H: Tom Glavine plays for the Atlanta Braves.

Atlanta Braves pitcher Tom ->

Tom plays for the Atlanta Braves.

There are no baseball teams in ResearchCyc. Also, *Pitcher-TheWord* has only one denotation, and it's a *ServingVessel*, not a hurler.

4. IE 268

T: There can be no doubt that the Administration already is weary of Aristide, a populist Roman Catholic priest who in December, 1990, won an overwhelming victory in Haiti's only democratic presidential election.

H : Aristide became president of Haiti in 1990.

won victory in presidential election -> became president

Win-TheWord has only one *verbSemTrans* in ResearchCyc, as a transitive verb that takes one NP argument. There aren't translations for the *DitransitiveNP-PPFrame*, which is how we parse this example.

We easily find that "victory" is the object of the verb "win"; ResearchCyc would assert given our dependency tree that Aristide is a *winner-First* of some NP headed by "election". However, that's not clearly connected with the notion of "president" in ResearchCyc.

5. CD 674

T: Jakarta lies on a low, flat alluvial plain with historically extensive swampy areas; the parts of the city farther inland are slightly higher.

H: The parts of Jakarta away from the coast are on slightly higher land.

farther inland ->

away from the coast

The only ResearchCyc predicate containing some notion of "farther" is *fartherNorthThan*, and it is not related to any more generic lexical entries.

6. CD 801

T: *Reagan was seriously wounded by a bullet fired by John Hinckley Jr.*

H: *John W. Hinckley Jr. shot Reagan in the chest.*

The ResearchCyc entry for *Wound-TheWord* denotes an *IncurringAnInjury* action. Unfortunately, the only *verbSemTrans* for *Shoot-TheWord* interprets the word as a denoting a *VisualImageRecording* action.

There is a *denotation* assertion stating that *ShootingAProjectileWeapon* is a concept denoted by the second sense of the verb *shoot*. However, there's no clear relationship, direct or indirect, between this predicate and *IncurringAnInjury*.

7. IR 102

T: *The White House failed to act on the domestic threat from al Qaeda prior to September 11, 2001.*

H: *White House ignored the threat of attack.*

Given the text sentence with correct dependencies, ResearchCyc would interpret “failed to act” in the *TransitiveInfinitiveVerbFrame* as denoting a *failureForAgents* relation between the subject “White House” and the action denoted by the INF-COMP “to act”. (There is a multi-word string entry for ResearchCyc for “fail to make a payment”, but none for “fail to act”).

Also, *Ignore-TheWord* has no lexical information in ResearchCyc.

8. IR 36

T: *Scripps Memorial Hospital Encinitas emergency room doctors and nurses treat two to three injured surfers.*

H: *Scripps Hospital assists surfing accident victims.*

We've also got some good ideas about scoring verb similarity in a more informed (or at least different) manner than WordNet. Once we've identified the head verbs here (irrelevant of the NPs), we find a *verbSemTrans* for each and look at the *isa* for the head keyword: here *assist-TheWord* denotes a *HelpingAnAgent* action, with a *beneficiary* and a *performedBy* slot, and “treat” denotes a *MedicalTreatmentEvent* action, a direct hypernym, roughly, in ResearchCyc talk of *ServiceEvent*, again a direct hypernym of *HelpingAnAgent*.

In addition to this *genl/genl* relationship, we see that both verbs have *performedBy* and *beneficiary* roles (two for two), indicating a not-entirely-superficial similarity that could be leveraged to score these verbs as close in meaning. One can see how these inferences could be easily generifed as well.

9. IR 52

T: *Phish disbands after a final concert in Vermont on Aug. 15.*

H: *Rock band Phish holds final concert in Vermont.*

There's no lexical entry in ResearchCyc that has an appropriate translation for *Hold-TheWord*, namely the only *verbSemTrans*' in ResearchCyc corresponding to "holding" are *HeldCaptive*, *HoldingWithHand*, and *HoldingAnObject*, which specifies a physical object as an argument.

There is a *holds-Underspecified* relation that specifies a generic "holding" relationship, but this appears too complicated to be treated as a case in general.

10. IR 64

T: *The wait time for a green card has risen from 21 months to 33 months in those same regions.*

H: *It takes longer to get a green card*

ResearchCyc has a good *semTrans* for *Rise-TheWord* denoting an *IncreaseEvent* action, where the *objectActedOn* is in the subject position. The construction "It takes longer" here is a tricky one, though; I don't see any way that ResearchCyc could interpret this usefully.

11. IR 79

T: *The privately owned spacecraft only got about 400 feet into space, according to radar measurements, but it was enough to confirm that it no longer takes a well-heeled government project to organize space travel.*

H: *private spaceship launches.*

The single *verbSemTrans* of *Get-TheWord* requires an ADJP complement, where it denotes an *IntrinsicStateChangeEvent*: the object of state change (the SUBJ) is the

argument of (*toState SUBJ ADJ*).

Into-TheWord has only *prepSemTrans* in the *VerbPhraseModifyingFrame*. It's not clear that we could reconcile these expected differences (*ADJP* complement vs. *VP* modifier). It's also the case that the only *verbSemTrans* for *Launch-TheWord* requires an *NP* complement.

12. MT 1228

T: *An official of Abyan police, where 16 Western tourists are being held since yesterday, announced that the hostages are held by the Yemeni "Islamic Jihad" group, which is demanding the release of its leader and lifting the embargo on Iraq.*

H: *The Yemen branch of the "Islamic Jihad" group, kidnapped the 16 Western tourists.*

hostages are held by the Yemeni group -> the Yemeni group kidnapped the tourists

Kidnap-TheWord is translated as a *KidnappingSomeone* action with a *perpetrator* slot.

KidnappingSomeone is a spec of *ActsCommonlyConsideredCriminal*, *TakingAPersonPrisoner*, and *CriminalAct*.

There is a *compoundString* entry for "hold hostages" that denotes a *HoldingHostages* action. *HoldingHostages* is a spec of *ActsCommonlyConsideredCriminal* and *HostileSocialAction*. Aside from knowing that both actions are criminal, *ResearchCyc* doesn't connect this with more specificity.

13. PP 487

T: *Located just three miles from Tullamore and only 45 minutes from the K Club, venue of the 2006 Ryder Cup, is Esker Hills, a genuine hidden gem and one of Irish golf's best kept secrets.*

H: *The K Club will host the 2006 Ryder Cup.*

K Club, venue of the Ryder Cup -> K Club will host the Ryder Cup

ResearchCyc provides us here a translation of "venue" to *eventOccursAt* with the correct filler slots. The translation of *host-TheWord* in the verb frame invokes a *hostOfEvent* action, though there's not a clear connection between this and *eventOccursAt*.

14. QA 1454

T: *In fact, Woolsey had had no first-hand experience with the world of spies until President Bill Clinton appointed him Director of Central Intelligence.*

H: James Woolsey is the director of the CIA.

The only relevant translation of *appoint-TheWord* invokes an *AppointingAmbassador* action. The assertions at this level are specific to “The collection of events where a state appoints an ambassador to another state” and can’t handle the sort of event analysis of a “director appointment” required to help on this entailment.

15. CD 693

T: *This growth proved short-lived, for a Swedish invasion (1655-56) devastated the flourishing city of Warsaw.*

H: *Warsaw was invaded by the Swedes in 1655, and the city was devastated.*

Swedish invasion devastated Warsaw -> Warsaw was invaded by Swedes

Identifying that the “city was devastated” is straightforward here: we have a direct match in the dependency parse. The tricky issue in this example is how to understand the first half of the hypothesis. ResearchCyc identifies that “invasion” is the singular form of *Invade-TheWord*, which denotes a *MilitaryInvasion*. Unfortunately, there isn’t any lexical information about how to translate *Invade* in ResearchCyc, for example a statement about a how an invasion may have a “performedBy” slot; *Invade-TheWord* currently has no lexical assertions at all. This hinders us from using ResearchCyc to our advantage in this entailment.

16. CD 735

T: *Even more than other economic activities, Mexico's financial services are concentrated in the capital.*

H: *Industry, retail stores, finance, and communications are all centered in the capital.*

Financial Services -> Industry, retail stores, finance, and communications

This example would require our system to do a particular sort of noun phrase matching order to identify each enumerated noun was a type of “Financial Services”. This sort of noun-phrase entailment is something that would be generally useful were it robust; our system already implements a ‘NP-match’ function which currently relies only on WordNet.

Such a function could conceivably query ResearchCyc to determine better-defined hypernymy in this case identifying that the hypothesis subject NP is composed of parts all of which are conceptual instances of a matching slot present in the text.

The only translation for the text NP here is the ResearchCyc assertion that the multi-word string “Financial Services” denotes a *FinancialCompany*, which in this case is not a hypernym of “finance” (which denotes *FinancialOrganization*) or “communications” (which doesn’t have an appropriate lexical entry, the only one has to do with *CommunicationEffectiveness*).

17. CD 767

T: *Hepburn, a four-time Academy Award winner, died last June in Connecticut at age 96.*

H: *Hepburn, who won four Oscars, died last June aged 96.*

ResearchCyc has got no entry for “Academy Award” or “Oscar”.

18. CD 779

T: *Voting for a new European Parliament has been clouded by apathy.*

H: *Apathy clouds EU voting.*

European Parliament : EU

ResearchCyc relates the concept *EuropeanUnion* to the string “EU” with the *initialismString* predicate, (a special case of *acronymString* where the string is formed using the first letters of the constituent words. *InitialismString* connects over 500 abbreviations to concepts in ResearchCyc). We could replace this string in the parse, and our similarity measure would here assign a higher score matching “European Union” to “European Parliament” than if we had only had “EU”.

19. CD 820

T: *Kessler’s team conducted 60,643 face-to-face interviews with adults in 14 countries.*

H: *Kessler’s team interviewed more than 60,000 adults in 14 countries.*

Conducted interviews = interviewed

We’ve got machinery in place to handle the “more than 60,000” = 60,643 equivalency.

What is crucially missing on our side is the bolded equivalence above. The only

appropriate lexicalization of the *Conduct-TheWord* would insist that “Kepler’s Team” was the *directingAgent* of some ACTION called “interview”. *Interview-TheWord* has an *agentiveNounSemTrans* that looks like this:

agentiveNounSemTrans Interview-TheWord 0 GenitiveFrame

(and

(interviewee ?ACT :POSSESSOR)

(interviewer ?ACT :NOUN)))

It’s not clear to me whether this is appropriate for a translation of the instance above.

20. IR 128

T: *Hippos do come into conflict with people quite often.*

H: *Hippopotamus attacks human.*

ResearchCyc actually has an entry for *Hippo-TheWord* linking the occurrence to the concept *Hippopotamus*. What remains then is to understand the “come into conflict” construction. The lexical information for *conflict-TheWord* is sparse, and there aren’t any *multiWordString* definitions that capture the meaning of this construction.

A Toy Example

We construct this simple but non-trivial example as a starting point to help us understand what machinery needs to be in place to get arguments aligned and translated to meaningfully related ResearchCyc concepts. It’s worth walking through how we’d set up a ResearchCyc query to ask about lexicalizations given typed-dependency output.

T: *John bought a car from Paul.*

F: *Paul sold a car to John.*

The Stanford parser gives us these typed dependencies from the sentences:

T: **nsubj**(bought, John) **det**(car, a) **dobj**(bought, car) **from**(bought, Paul)

F: **nsubj**(sold, Paul) **det**(car, a) **dobj**(sold, car) **to**(sold, John)

Our system identifies the head verb of each sentence and attempts to find the best fitting

subcatFrame for each given the dependency parse. In this case, we see that the head verb in both cases has two complements, and NP and a PP.

Starting with the Text sentence, we check to see if there is a *DitransitivePP-NP* frame for *Buy-TheWord*. There is not. Then we look to see if there is an equivalent translation, in this case using (*PPCompFrameFn DitransitivePPFrameType From-TheWord*) to represent the subcat frame where there is an NP and a PP argument headed by “from”. It’s my understanding that this redundancy exists because the *DitransitivePP-NP* captures translations that are valid independent of the preposition, whereas the latter places constraints on the preposition and correspondingly the semantic translation. The ResearchCyc translation tells us that the following are true:

(*isa :ACTION Buying*)
(*seller :ACTION :SUBJECT*)
(*objectPaidFor :ACTION :OBJECT*)
(*buyingPerformer :ACTION :OBLIQUE-OBJECT*)))

On to the hypothesis,: we check to see if there is a *DitransitivePP-NP* frame, there is not. We check to the *PPCompFrame* as above and find the following translation:

(*verbSemTrans Sell-TheWord 3*
(*PPCompFrameFn DitransitivePPFrameType To-TheWord*)
(*and*
(*isa :ACTION OfferingForSale*)
(*performedBy :ACTION :SUBJECT*)
(*transferredObject :ACTION :OBJECT*)
(*target :ACTION :OBLIQUE-OBJECT*)))

In fact this is a problem. We would have got the correct translation (which corresponds to a *Buying* action, not an *OfferingForSale* action) only if we had asked for the *TransitiveNPFrame*. Unless we explicitly tried both, we’ve introduced a mistranslation which isn’t necessarily recoverable, because the relationship between *OfferingForSale* and

Buying isn't well-defined in ResearchCyc.

Conclusions

We present some summary remarks on the utility of ResearchCyc in recognizing textual entailment using our NL.

Lexical Coverage

To be sure, lexical coverage is the deficiency in ResearchCyc which hurts us the most on this task, and it is especially problematic in the absence of functional ResearchCyc NL tools. In most cases we find sparse or suboptimal lexicalizations that render any further search useless. Even on our toy example, the absence of a proper translation for “sells *X* to *Y*” keeps us from making the meaningful connection that we would expect from ResearchCyc: that both verbs express a buying action and can be translated as such given their NP-PP arguments.

True, we can implement searches that traverse the space of ResearchCyc relations and probably get some utility even if we have mistranslated the verb, but we would hope that for most examples that the right translation is in the KB: even too many ambiguous translations would be better than none.

Concept Linkage

It is hard to discuss this in general given the expert nature of much of ResearchCyc's knowledge, but for our purposes the concept linkage is also lacking in most examples: empirically speaking, we can almost never get from one sentence to the other using ResearchCyc alone. To this extent, ResearchCyc as a standalone RTE system is currently infeasible. Word-level similarity modules (that tell us that “hippo” means “hippopotamus”, or that a “mosque” is a “building”, or that “EU” designates the European Union), however, may be generally useful even in the situation where ResearchCyc can't handle arbitrary lexical lookups and conceptual connections. We intend to further explore using ResearchCyc for such similarity calculations in future work.