

707.000
Web Science and Web Technology
„User Intentions and Intentional Structures
on the Web“

Markus Strohmaier

Univ. Ass. / Assistant Professor
Knowledge Management Institute
Graz University of Technology, Austria

e-mail: markus.strohmaier@tugraz.at
web: <http://www.kmi.tugraz.at/staff/markus>

Overview

Home Assignment 6

- Request for Deadline Extension
- Extended Deadline: FR 6.6.2008, 23:59
 - Delayed submission comes with a penalty: You can only receive 3/5 of the total points of HA6. This amounts to a maximum of 15% instead of 25% for HA6
 - If you re-submit your previously submitted version after today's deadline,
 - we will evaluate the latest submission
 - the „delayed submission“ penalty will apply

Overview

Agenda

- **History of smileys :-)**

A glimpse at recent and ongoing research at the KMI:

- [How do users express goals on the web? - An exploration of intentional structures in web search.pdf](#)
- [Different degrees of explicitness in intentional artifacts - Studying user goals in a large search query log.pdf](#)



17-Sep-82 14:59 Joseph Ginder at CMU-10A (*%)
 I believe that the joke character should be % rather than *.
 =

Joke

16-
 19-Sep-82 18:56 Jeff Shrager at CMU-10A 38521,03,9(6),9(9),1(5),0
 Just signifying that a message is a joke is certainly not sufficient.
 One can develop a taxonomy of bboard message types along several different
 dimensions. Also, where a continuum is preferable to a taxonomy (such as
 where humor value is at issue) one can similarly use a scale to indicate
 where along that scale this message lies. Suppose that all dimensions are
 referred to by a ten point scale (we'll use all integers here although one
 can certainly imagine reals in the case of fine grain continuous scales).
 Some dimensions will be bitwise encoded as well.
 Here is a sample of a coding scheme:

Reply

* Proposal

*%/&/#
 Proposals

COMMUNITY: (this is a binary scale with a bit position for
 each department totalling about 32 bits)
 TOPIC: (two digits 00-99)
 (00) Political, (01) Scientific, (02) Computer, (03) Meta, etc
 FLAME VALUE: (continuous 0.0-10.0)
 HUMOR VALUE: (0.0-10.0)
 BORDOM VALUE: (0.0-10.0)
 INFORMATIONAL CONTENT: (-10.0 (for queries) to 10.0 (for their answers))

:-)
 Proposals

Quantitative
 Proposals

Note that some of these scales are purely according to the opinion
 of the author. Thus, we provide, also, a confidence scale: to go along
 with each continuous scale (to be enclosed in parens after the value).
 =
 the joke.)

“How culture made your modern mind”

17 May 2008, New Scientist Print Edition, Andy Coghlan

The Theory of mind:

- denotes a specific cognitive capacity
 - the realisation that other creatures are intelligent and capable of independent thought and intention

- This capacity was a critical neurological breakthrough

It enabled humans to cooperate on tasks such as using the combined strength of several people to move a heavy object or hunt together.”

Günther Knoblich, University of Birmingham, UK,
[New Scientist, May 17, 2008]

- **We have largely lost this capacity on the web**

Informally:

“the ability to attribute mental states—beliefs, intents, desires, pretending, knowledge, etc.—to oneself and others and to understand that others have beliefs, desires and intentions that are different from one's own.” [WIKIPEDIA]

“Intentional Translucence”

- Social Translucence [Erickson2003]
 - In the digital world we are socially blind
 - Making social information visible on the web
 - Users being aware of what is happening
 - Users being held accountable for their actions as a consequence of that awareness.
 - Privacy vs. Visibility
- **Intentional Translucence**
 - In the digital world, **we are blind** with regard to the intentions of others
 - **Making intentional information visible** on the web
 - Enabling users to be aware of the intentions of others
 - Enabling users to cooperate on achieving goals
 - Understanding „traces of intent“ and clues related to intentions
 - Privacy vs. Visibility



Figure 1.4. The Queue proxy

go skiing find books Understand Mozilla
 information about tugraz find a Job in Graz
 relax understand linear algebra create Fire
 know SIMBA order asian take-out
 organise my bookmarks find

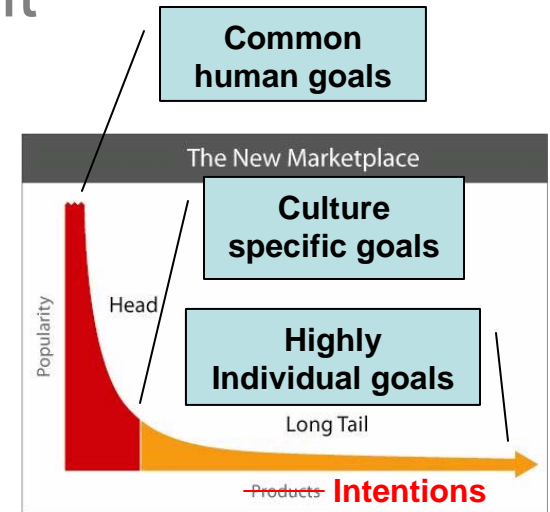
From Content to Intent

The long tail [Anderson]

- Huge diversity of user goals on the web, most are implicit
- majority of goals are “unpopular”, i.e. shared by only a few
- Yet, on a global scale, there is community of people who share highly individual goals e.g. “build an english cottage”

But the web makes it difficult for users

- To be aware of the intentions of others
- To allow users to collaboratively achieve goals



[Anderson 200X]

From Content to Intent

- Goal acquisition and –mining
 - harvesting “traces of intent”
- Goal representation and modeling
 - identifying relations between goals
- Social goal discovery and intentional translucence
 - goals as “first class citizens”

Nr.	Query	Time Stamp
#1	Entertainment tv	2006-03-30

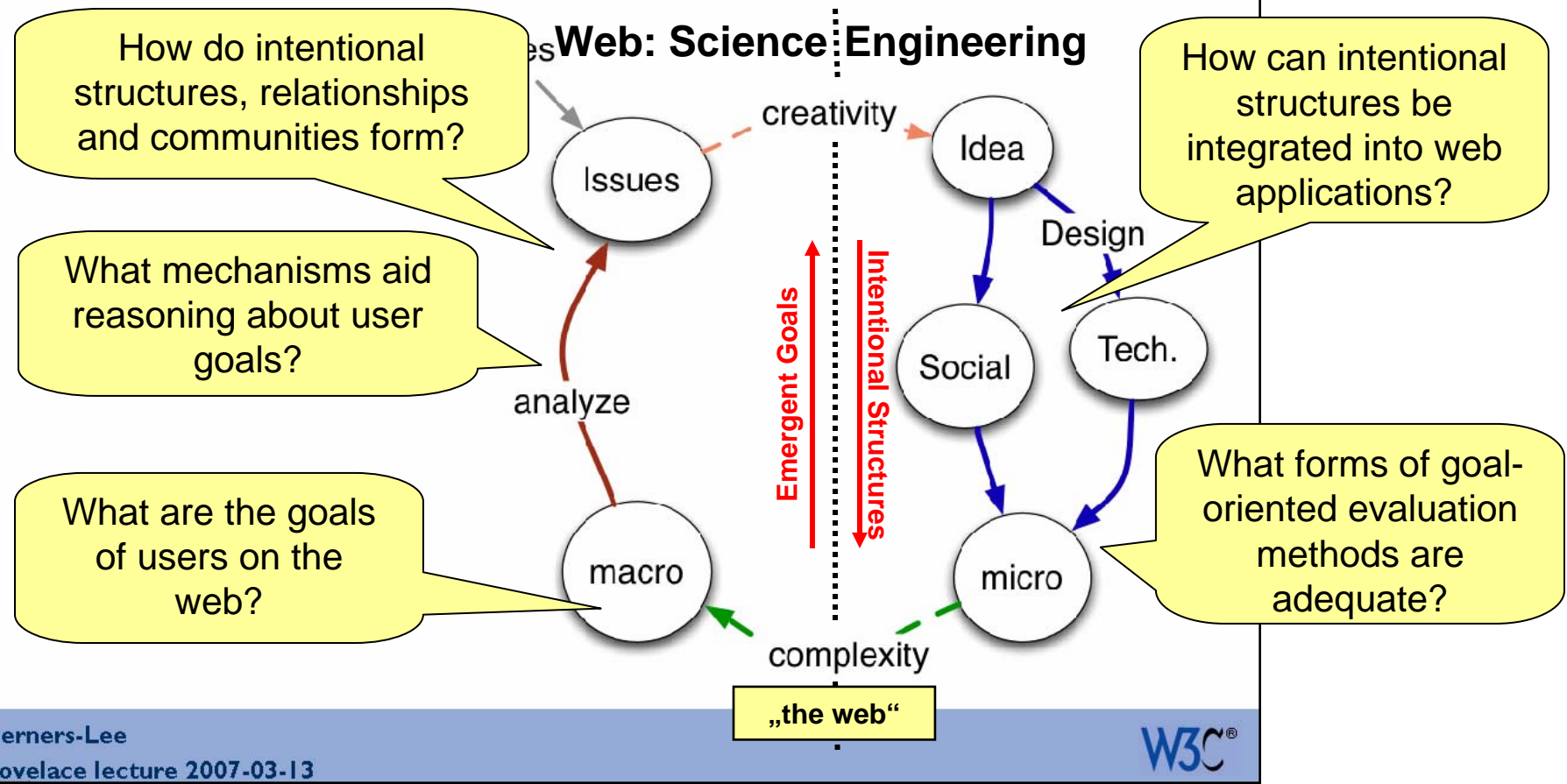
Skydive laugh more

go skiing find books Understand Mozilla
 information about tugraz find a Job in Graz
 relax understand linear algebra create Fire
 know SIMBA order asian take-ou
 organise my bookmarks find

An *Intentional* Web

Goal-Oriented Analysis and Engineering of Social-focused Web Applications

The process of web science



Berners-Lee
Lovelace lecture 2007-03-13



Motivation

Search engine vendors are desperate to learn more about users intentions

Peter Norvig, Director of **Google** Research

- understanding users' needs to a greater extent represents an “*outstanding*” research problem.
- Google is currently looking at “*finding ways to get the user more involved, to have them tell us more of what they want.*” [2]

Yahoo's future strategy aims at “driving deeper engagement” with its users through vertical search, personalization, desktop search and social search and believes the increasing sophistication of user search behavior supports the goal (2006)

<http://blogs.zdnet.com/micro-markets/index.php?p=27>

-The Database of Intentions-

John Battelle

- The „Database of Intentions“ is

“The aggregate results of every search ever entered, every result list ever tendered, and every path taken as a result. [...]

This information represents [...] a place holder for the intentions of humankind - a massive database of desires, needs, wants, and likes that can be discovered, subpoenaed, archived, tracked, and exploited to all sorts of ends.

Such a beast has never before existed in the history of culture [...].”

Source: <http://battellemedia.com/archives/000063.php>, last accessed on July 3rd, 2007

Selected Research Questions

1. How can we **identify goals** in search query logs?
2. How can we represent search goals in a **semi-formal goal graph**?
3. How can we represent the search process as a **traverse** through the goal graph?
4. How can goals **direct search behaviour**?
5. Is there a difference between **explicit and implicit intentional search queries**, and how can it be identified?
6. How can we construct **large scale goal graphs**?

Example: AOL Search Database

- History
 - Released with the purpose of serving science
 - More than 20 million queries
 - User IDs, timestamps, queries, URLs selected
- Implications
 - “A Face Is Exposed for AOL Searcher No. 4417749”
<http://www.nytimes.com/2006/08/09/technology/09aol.html?ex=1312776000&en=f6f61949c6da4d38ei=5090>
 - „57thiest dumbest moment in business“
according to Business 2.0 Magazine on CNNMoney:
http://money.cnn.com/magazines/business2/101dumbest/2007/full_list/index.html
- Data online
 - http://sergiorebelo.com/twodotfive/?page_id=25
 - Links to several online services for searching the AOL search database (September 2007)
 - <http://www.aolstalker.com/>
 - „Überwacht die Überwacher“

Motivation

Today,

- many activities on the web are *driven by goals* (e.g. *buy car, find gift, treat influenza*)
- many services and resources on the web already provide varying support for *achieving goals*
- Users routinely *translate high-level goals* into specific search queries, tag concepts or classification terms

yet

- The web lacks explicit **intentional (i.e. goal-oriented) structures**

Which makes it

- hard for users to explore and express goals
- difficult for technologies to discover, represent, comprehend and reason about users' goals
- difficult for communities to share and discover goals
- ...

Goals and Intentions

Goals are
„desired states of affairs“

Intentions are
„goals + commitment“
[Mylopoulos 2006]

Some high-level examples of potential users' goals on the web:

- Gifts -> create | find | buy | sell | ...
- Wedding -> organize | attend | sabotage | film | ...
- Car -> buy | sell | repair | tune-up | insure | ...
- Influenza -> treat | avoid | eliminate | ...

Structure of Goals

(in Requirements Engineering)

Simplified Assumption:

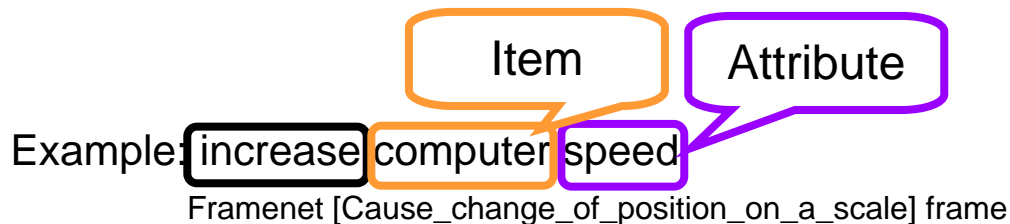
Goal = Verb + Semantic Frame elements
(see, e.g. [Liaskos 2006])

Example verbs: Achieve, Make, Improve, Speedup, Increase, Satisfied, Completed, Allocated, Maintain, Keep, Ensure, Avoid, Know, Monitor, Track, Provide, Supply, ... [Regev 2005]

The Berkeley Framenet Project [Ruppenhofer et al 2006]:

Semantic Frame: “a script-like conceptual structure that describes a particular type of situation, object or event along with its participants and props”

Frame Elements: Each of these elements of a semantic frame is called frame elements



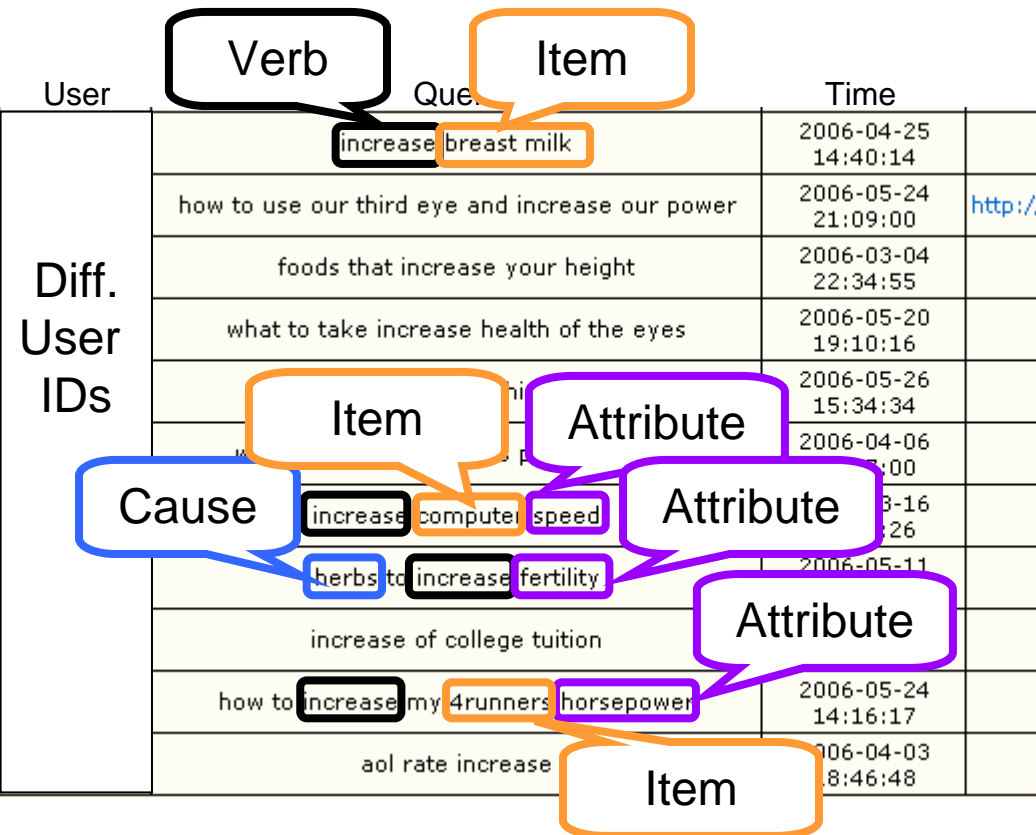
Example: Occurences of „Increase“ in Search Queries

Query for verb „increase“, excerpt of search results, annotated with frame elements

User	Query	Time	
Diff. User IDs	increase breast milk	2006-04-25 14:40:14	
	how to use our third eye and increase our power	2006-05-24 21:09:00	http://
	foods that increase your height	2006-03-04 22:34:55	
	what to take increase health of the eyes	2006-05-20 19:10:16	
	increase hi	2006-05-26 15:34:34	
	increase computer speed	2006-04-06 17:00:00	http://www.ehow.com
	herbs to increase fertility	2006-05-11 13:16:26	
	increase of college tuition		
	how to increase my 4runners horsepower	2006-05-24 14:16:17	
	aol rate increase	2006-04-03 18:46:48	http://www.forbes.com

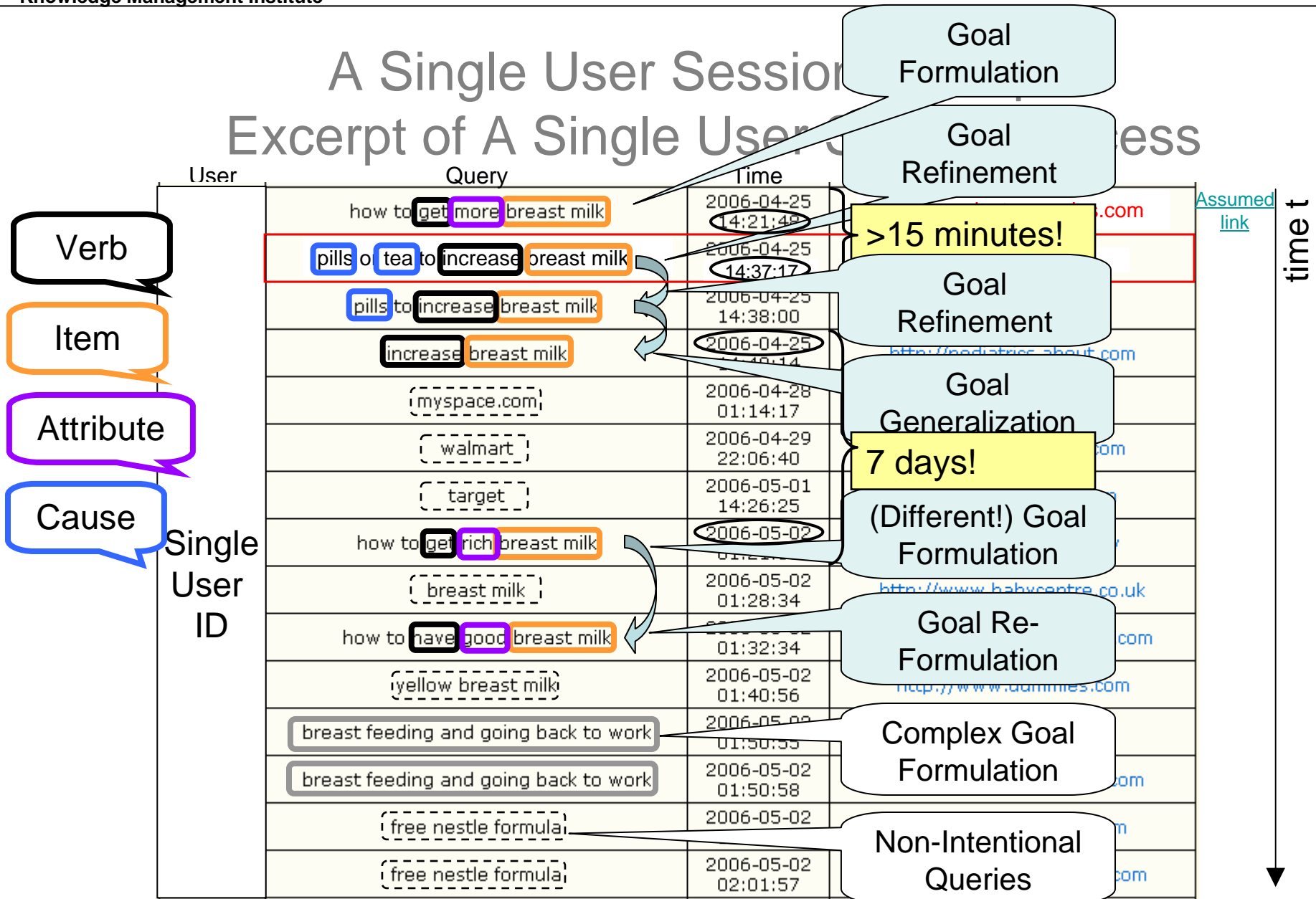
Core Frame Elements for [Cause_change_of_position_on_a_scale]

- Agent
- Attribute
- Cause
- Item



Source: AOL Search database, retrieved from <http://www.aolsearchdatabase.com/>, accessed on July 2nd, 2007

A Single User Session Excerpt of A Single User Session



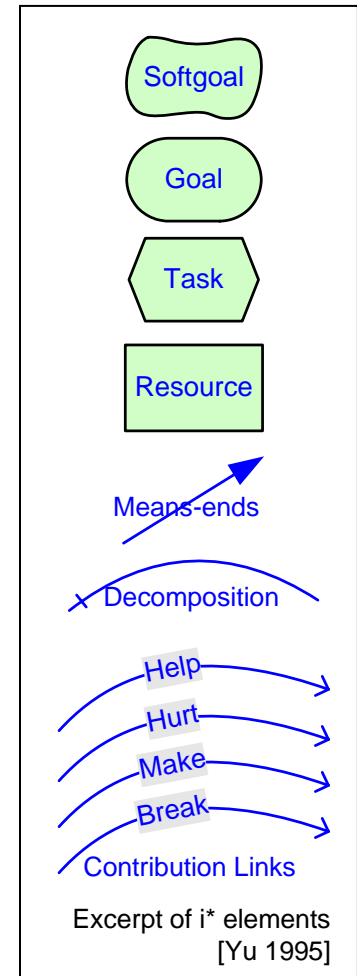
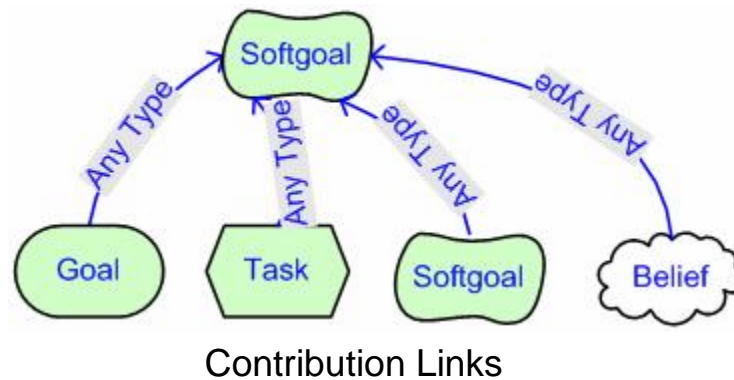
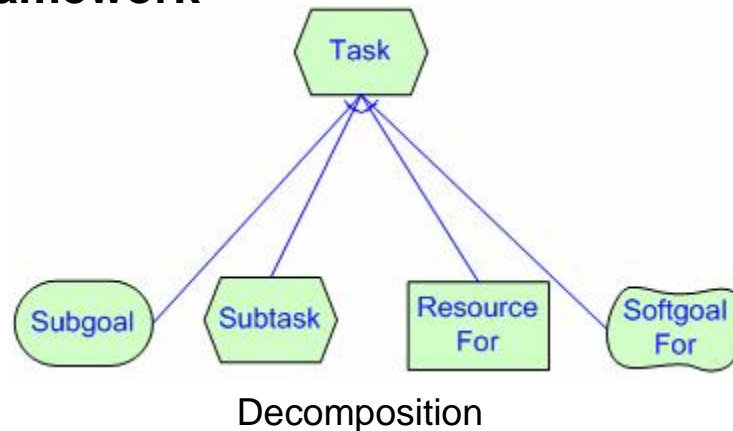
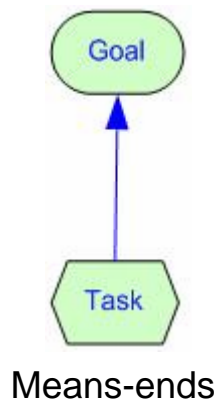
Questions

1. How can we represent the involved goals in a **semi-formal goal graph**?
2. How can we represent the search process as a **traverse** through the goal graph?

Intentional Structures - Goal Models in Requirements Engineering

Excerpts of the i* framework

[Yu 1995]



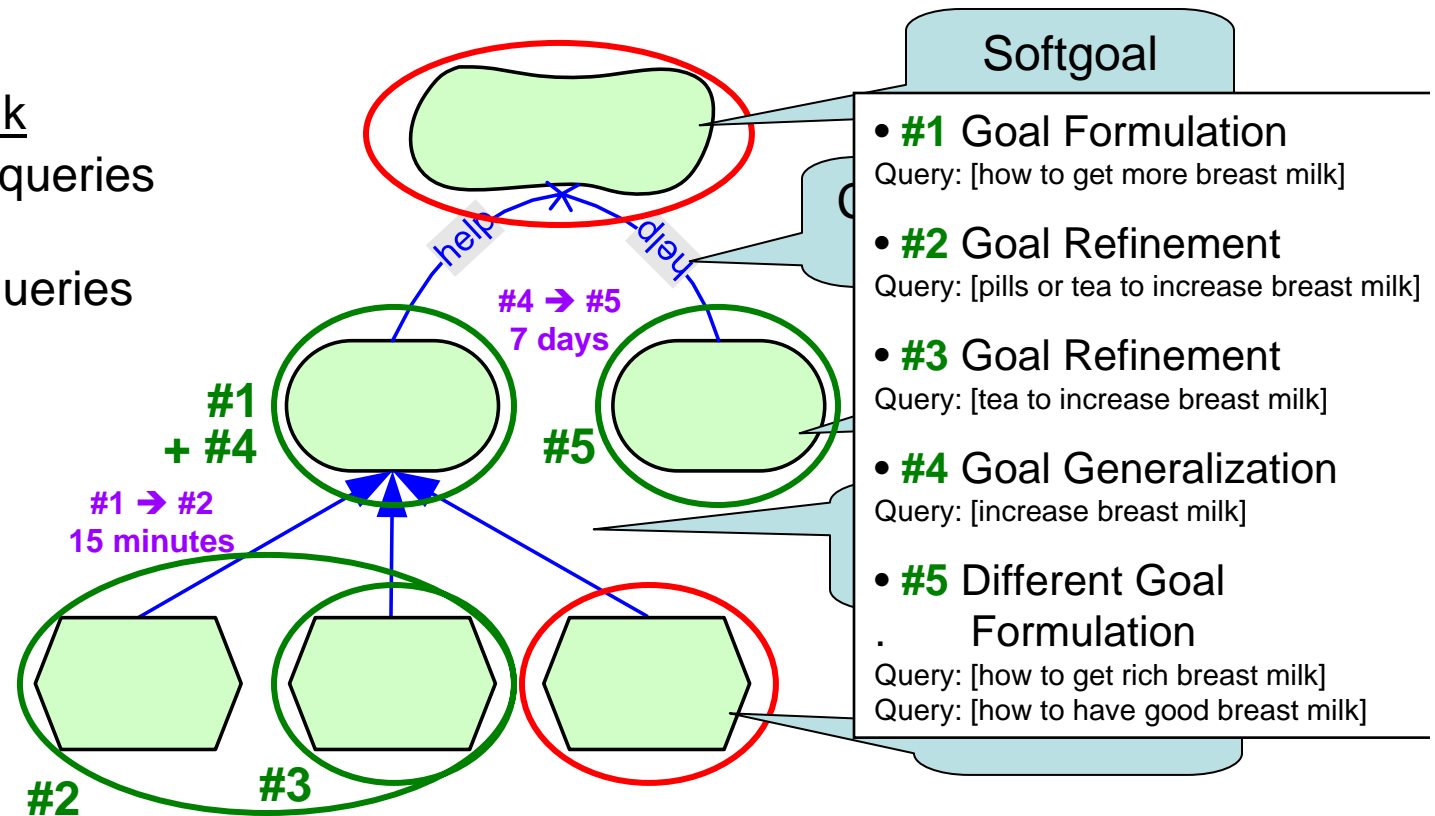
Source, <http://istar.rwth-aachen.de/tiki-index.php?page=iStarQuickGuide>, last accessed on July 3rd, 12007

Questions

1. How can we represent the involved goals in a **semi-formal goal graph**?
2. How can we represent the search process as a **navigation** through the goal graph?

Domain: Breast Milk

- Executed search queries
- Time differences between search queries
- Not-executed search queries



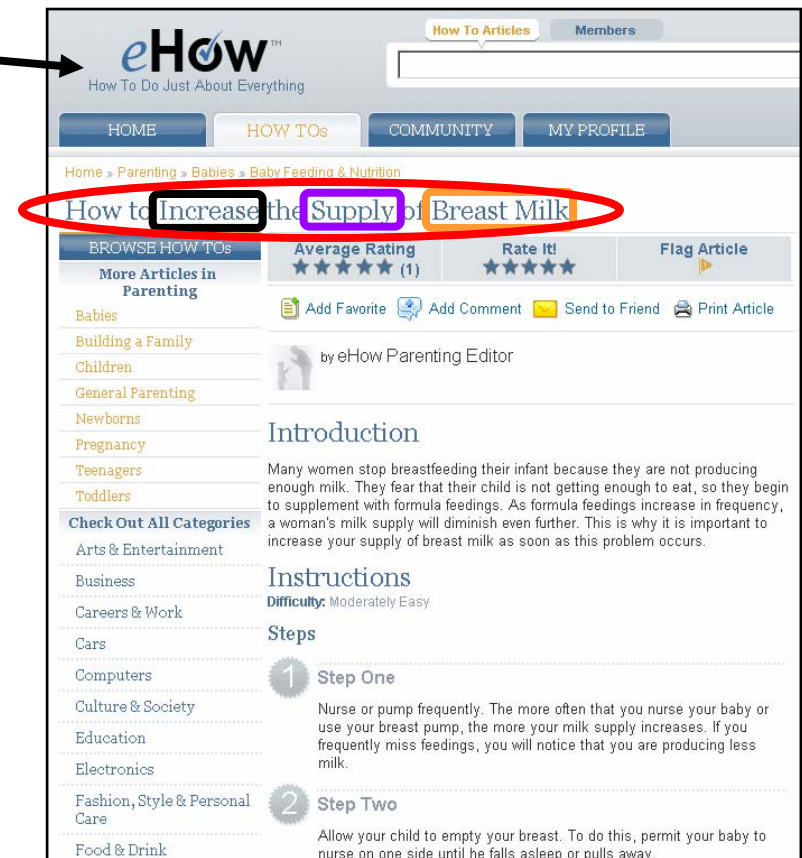
Based on the NFR/i* Framework

[Chung et al 2000/Yu 1995]

Other web2.0 applications „How to X?“

Examples:

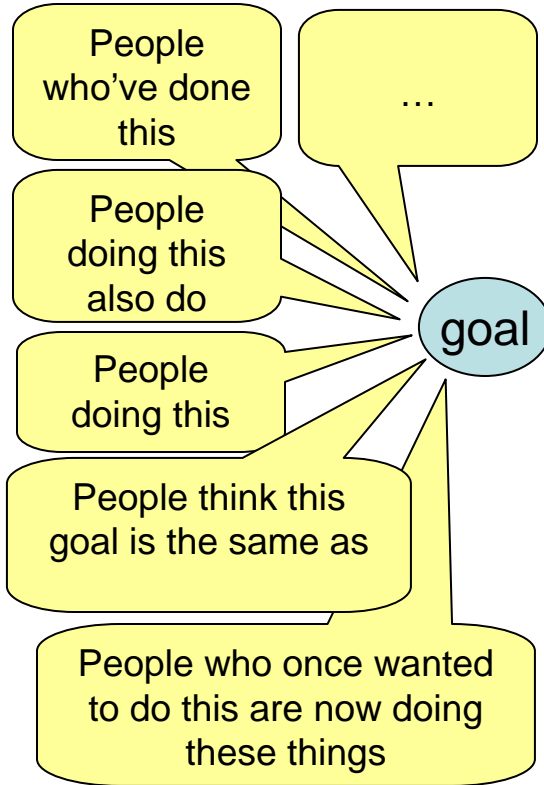
- [www.ehow.com /](http://www.ehow.com/)
www.wehow.com
how to descriptions & videos
- www.wikihow.com
- www.videojug.com
explanatory videos
- www.viewdo.com
includes “Required Tools” and “Required Parts”
- www.howto.tv
- www.43things.com
- More at
http://dir.yahoo.com/Reference/How_To_Guides/



43 Things API

Common XML Objects

- [Person](#)
- [Goal](#)
- [Entry](#)
- [City](#)



Goal Methods:

- [search_goals](#)
- [get_goal_by_id](#)
- [get_goal_by_name](#)
- [get_goals_similarities](#)
- [get_goals_people](#)
- [get_goals_entries](#)

People Methods:

- [search_people](#)
- [search_people_by_email](#)
- [get_person](#)
- [get_persons_completed_things](#)
- [get_persons_entries](#)
- [get_persons_progress_on_goal](#)
- [get_persons_teammates](#)
- [get_persons_neighbors](#)
- [get_persons_tags](#)
- [get_persons_tag_cloud](#)
- [get_subscriptions](#)

Team Methods:

- [get_teams_progress](#)

Entry Methods:

- [get_entry](#)

Tag Methods:

- [search_tags](#)
- [get_tags_goals](#)
- [get_tags_similarities](#)

City Methods:

- [search_cities](#)
- [get_city](#)
- [get_citys_people](#)

Action Methods

- [add_goal_by_id](#) (requires authentication)
- [add_goal_by_name](#) (requires authentication)
- [complete_goal](#) (requires authentication)
- [remove_goal](#) (requires authentication)
- [update_goal](#) (requires authentication)

ATOM Methods:

- [introspection](#)
- [entry](#) (requires authentication)

+ RSS

Source: http://www.43things.com/about/view/web_service_api,
last accessed on July 4th, 2007

Selected Research Questions

1. How can we **identify goals** in search query logs?
2. How can we represent search goals in a **semi-formal goal graph**?
3. How can we represent the search process as a **traverse** through the goal graph?
4. How can goals **direct search behaviour**?
5. Is there a difference between **explicit and implicit intentional search queries**, and how can it be identified?
6. How can we construct **large scale goal graphs**?

Result of one of the projects conducted by
„Knowledge Technologies“ students
C. Ruggenthaler und A. Haselsberger

In which principal ways can we modify search?

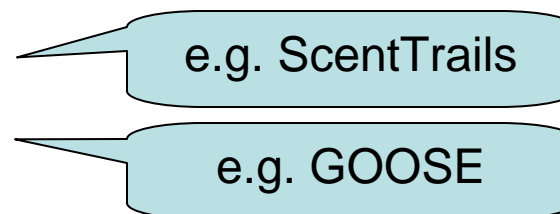
Query modification

Velez et al 1997

- **Scope reduction** (of underspecified queries)
(e.g. refinement, conjunction)
- **Scope expansion** (of overspecified queries)
(e.g. generalization, disjunction)

Result adaptation

- *Adapting the content*
- *Re-ordering the links*
- *Augmenting the links*



What can be adapted in adaptive hypermedia?
<http://wwwis.win.tue.nl/ah94/Brusilovsky.html>

Query Modification

Query Refinement is the incremental process of transforming a query into a new query that more accurately reflects the user's information need. [Velez et al 1997]

- Current implementations e.g. search for “car” on ask.com ->
 - new car, used cars, fast cars, sports cars, car insurance
 - predominately non-intentional

Our idea: **Intentional query expansion** focuses on query expansion exclusively from an intentional perspective:

- E.g. search for “wine crop” ->
 - **verbs:** protect wine crop, increase wine crop, improve wine crop
- E.g. search for “increase wine crop” ->
 - **means:** fertilizer to increase wine crop, insecticides to increase wine crop
 - **helps:** winery be successful, harvest more wine
 - **hurts:** produce organic wine

→ Support the user in formulating/refining/disambiguating/learning about his/her goals.

I) What is Intentional Query Expansion?

*“The purpose of query expansion in information retrieval is to make the user query **resemble** more closely **the documents** it is expected to retrieve”*

[Natural Language Information Retrieval: Trec-5 Report (1996) Tomek Strzalkowski, Louise Guthrie, Jussi Karlgren, Jim Leistensnider, Fang Lin, Jose Perez-Carballo, Troy Straszheim, Jin Wang, Jon Wilding, Text REtrieval Conference]

In Intentional Query Expansion,

*The purpose of query expansion is to make the user query **resemble** more closely **the intentions** a user has.*

- Users might not have clear expectations regarding the documents in the result set
- Users might not even have a clear understanding of the intentions they have (Open problems)

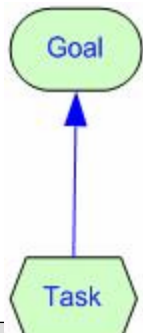
I) What is Intentional Query Expansion?

Definition:

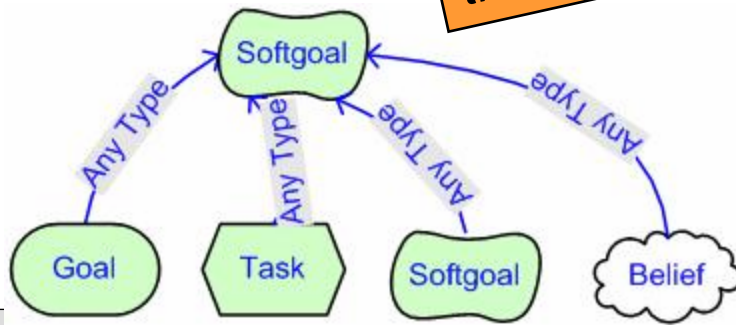
Intentional query expansion is the incremental process of transforming a query into a new query based on **intentional structures found in a given domain**

What are intentional structures?

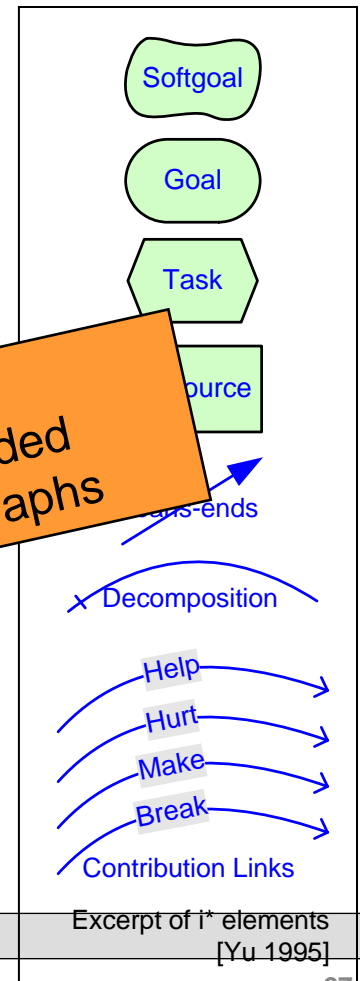
Intentional query expansion as a guided traversal of goal graphs



Means-ends

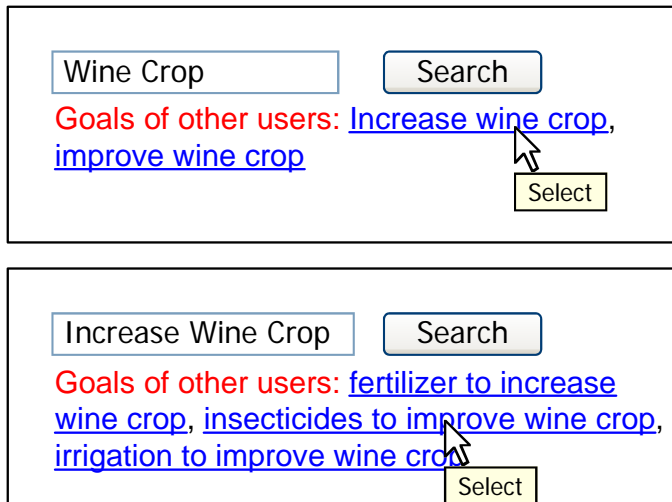


Contribution Links



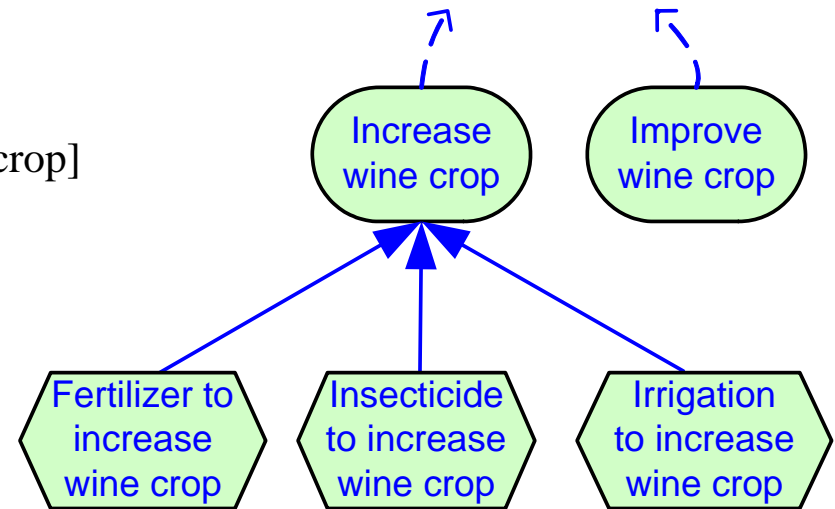
Excerpt of it* elements [Yu 1995]

II) Is Intentional Query Expansion Feasible? A Proof of Concept



increase [item wine crop]

[cause insecticide]



A Mockup of Utilizing Collaborative Intentional Structures to Facilitate Search on the Web

How can goals direct search behaviour? Experimentation

- Let's „buy a car“!
- Experimental Demonstrator:
<http://webdev.know-center.tugraz.at:8080/IQE3/>
[joint work with students: A. Haselsberger, C. Ruggenthaler, M. Rettenbacher]

An Exploratory Approach:

- Contacting several online resources just-in-time, such as *Yahoo Related Queries* and *43things.com Goals and Tags*
- Mining goals from identified text via POS analysis
- Structuring resulting set of goals
- Suggesting goals through **query expansions**



Project Searchengine: WT 07/08

buy car	buy cars	improve credit score	sell car	make money	get shape	get license	get job
find job	f--k car	complete education	clean credit	buy ?car	buy kia	buy enterprise	buy days
buy condominium unit	buy car online						

Search finished in 19017 ms

Current Page: 1

[Previous](#) [\[1\]](#) [\[2\]](#) [\[3\]](#) [\[4\]](#) [\[5\]](#) [\[6\]](#) [\[7\]](#) [\[8\]](#) [\[9\]](#) [\[10\]](#) [Next](#)

Total Search Results: 571000000

[Buy A Car Wash | Target.com](#)
 Shop, find, buy "Buy A Car Wash" and other related products at Target.com.
<http://www.target.com/gp/search.html?field-keywords=buy-a-car-wash>

[Buy Cars Online, New & Used Car Search, Auto Classifieds](#)
 New car prices, used cars for sale & car buying tips. ... New-Car Buying Guides. Compare Cars Side-by-Side. Compare Car Photos. Latest Rebates & Incentives ...
<http://www.cars.com/go/buyIndex.jsp>

[Buy a New Car on Yahoo! Autos](#)
 Get FREE price quotes on new cars from dealers near you. ... This service is totally FREE, with no obligation to buy. Select a Car and Enter Your Zip Code ...
<http://autos.yahoo.com/newcars/buy.html?refsrc=autos/home>

[Buying a Car](#)
 ... buy a used car from a private party directly, or a new or used car from a car dealer. ... Cars bought from an auto dealer already have inspection stickers ...
http://www.rpi.edu/dept/doso/ISSS/public_html/carbuyng.html

[Buying A Car Seat - Shopzilla.com](#)