Automating Web-Navigation Support Using a Cognitive Model

4th International Conference on
Web Intelligence, Mining and Semantics
June 2 , 2014 -Thessaloniki, Greece

Herre Van Oostendorp

Universiteit Utrecht

Sonal Aggarwal, Bipin Indurkhya

IIIT-H

Web-Navigation

- Good UI facilitates easy navigation
- Leads to less mouse-clicks & less **cognitive load**

- No standardization in UI
- Unclear menus & navigation structures

Ways for providing Navigation Support

- Recording browsing steps & allowing users to replay
- Representing navigational objects in form of a map
- Keeping track of user's behavior & navigation history to find out what could be interesting to the user (requires initial training of the system)
- Merging browsing & searching techniques, & using information foraging theory (based on information scent, the most relevant links on the web page get highlighted, leading to user goal)

Why a Cognitive Model for Web-Navigation simulation is preferred

- Predicts navigation pattern & provides support to user
- Works well with first time website visit, no initial training or browsing history required

LICAI (Linked Model of Comprehension-based Action Planning & Instruction taking)

SNIF-ACT (Scent-based Navigation and Information Foraging in ACT Architecture)

MESA (Method for Evaluating Site Architectures)

CoLiDeS (Comprehension-based Linked Model of Deliberate Search)

CoLiDeS

- The model uses Latent Semantic Analysis (LSA) to determine *information scent* between user goal & content of hyperlinks on a given web page
- Information scent - measure of value/cost/access path of information sources obtained from proximal cues
- LSA - An unsupervised machine learning technique that builds the semantic space;

  Includes computation of **term frequency, inverse document frequency (TF-IDF);**

  Cosine similarity is computed using **dot products** of the two TF-IDF vectors (one for user goal & the other for hyperlink-text)

Variants of CoLiDeS

- **CoLiDeS+** includes structural dimension: Path adequacy (semantic similarity between the navigation path traversed so far during the session & user's goal)

- **CoLiDeS+Pic** includes semantics from pictures (obtained through user annotation)

- **Proposed model** - CoLiDeS++Pic

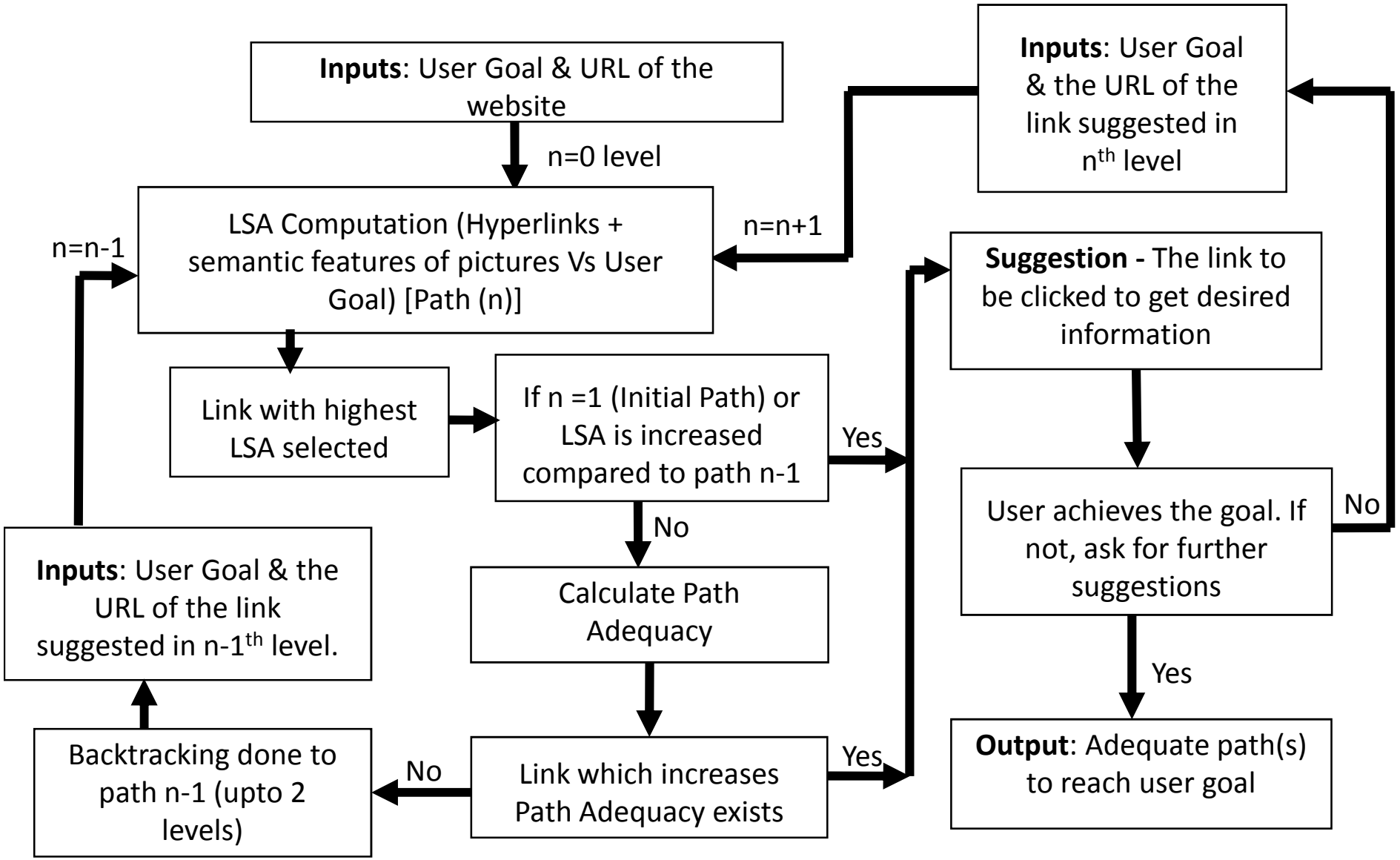
  Takes into consideration both structural dimension & information from semantic features of pictures.

  Expectations:

  i) CoLiDeS++PIC model will predict behavior better;

  ii) Information seeking performance will be improved with CoLiDeS++Pic

Process Diagram of the tool



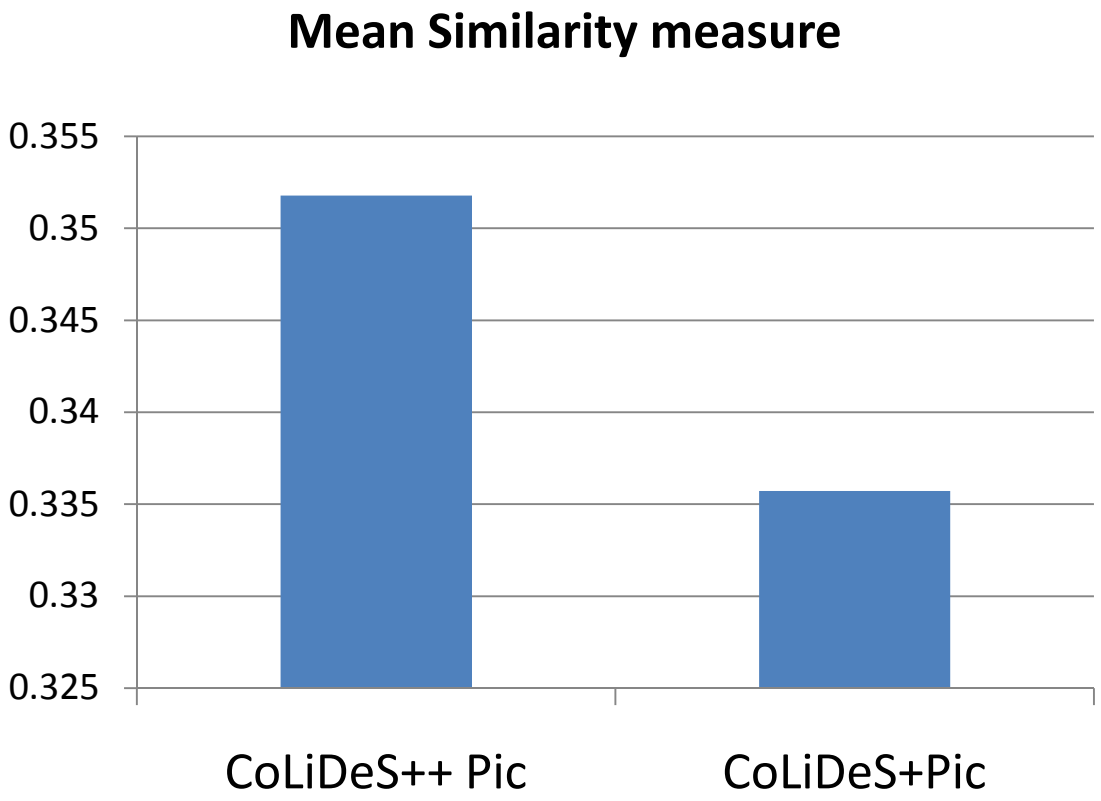
Example on a mock-up website

Level 1	LSA Value	
Respiratory System	0.251	
Nervous System	0.251	
Digestive System	0.270	
Circulatory System	**0.273**	
Level 2	LSA Value	Path Adequacy
Cardiovascular System	0.238	0.284
Lymphatic System	0.242	**0.308**

Note: bold links are selected by the system

User goal - "*Lymphatic System contains immune cells called lymphocytes, which protect our body from antigens. They are produced by lymph nodes. Name at least three locations in the body where lymph nodes are present*".

Model Behavior of CoLiDeS++Pic compared with CoLiDeS+Pic



Mean LSA values for correct links given by both models for 8 tasks on a given mockup website

Usefulness of the tool

Helps reducing cognitive load on the user & saves time

Particularly for:

- Visually-impaired persons
- Elderly people having memory problems
- New internet users
- Expert users doing multi-tasking

Validation Experiment

Goal: tool-support based on the links chosen by the model will help users during navigation;

particularly during multi-tasking.

Study on the influence of **support** and **multi-tasking** on performance of participants

(completion time, task accuracy, disorientation)

Validation Experiment

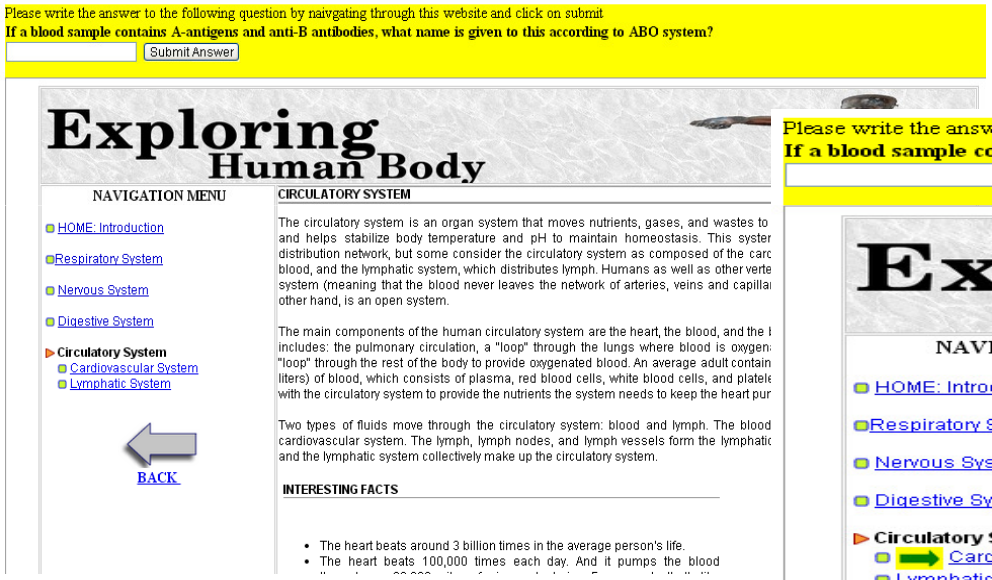
- 2x2x4 factorial design, where the first two factors (**tool-support and multi-tasking**) were between-subjects-variables, & the third factor (4 **levels of task**) was within-subjects.

- **Parameters measured** :
  Time, Accuracy, Disorientation (based on the ratio of the visited and the optimal node counts: L =√ ((N/ S - 1)2 + (R /N - 1)2) and R is minimum no. of pages needed to visit in order to finish task, S is the actual no. of pages visited, N is the no. of distinct pages visited & L is the disorientation. As users become increasingly lost the value of N/S tends to be 0 and also R/N tends to be 0, while for users who are not lost the ratios tend to be 1)

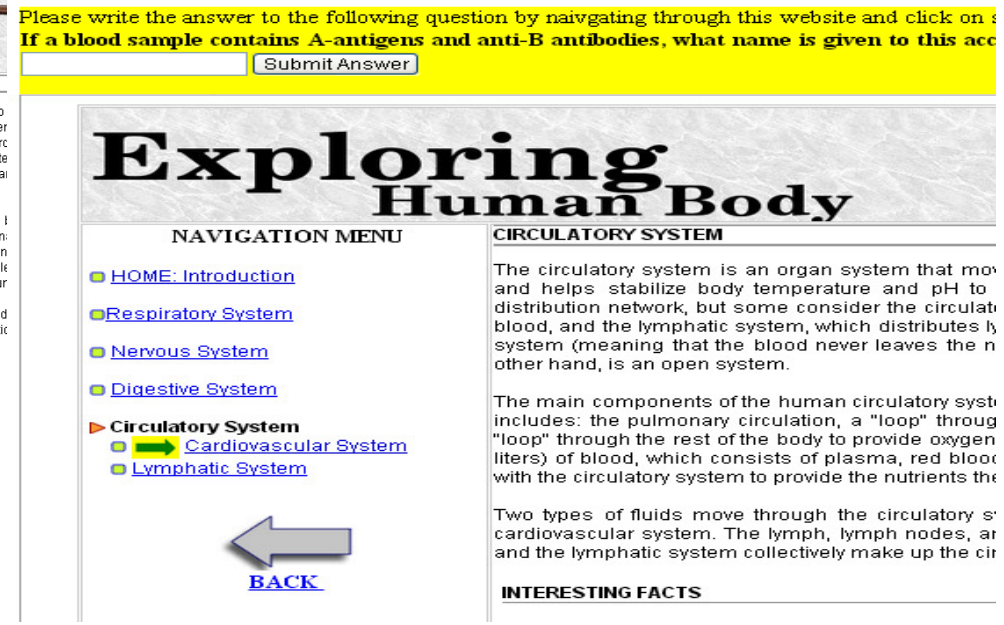
  Participants: 40 students of IIIT-H, 34M, 6F, (M age=27.14, SD=6.75).

Validation Experiment

- Mockup website (two versions)



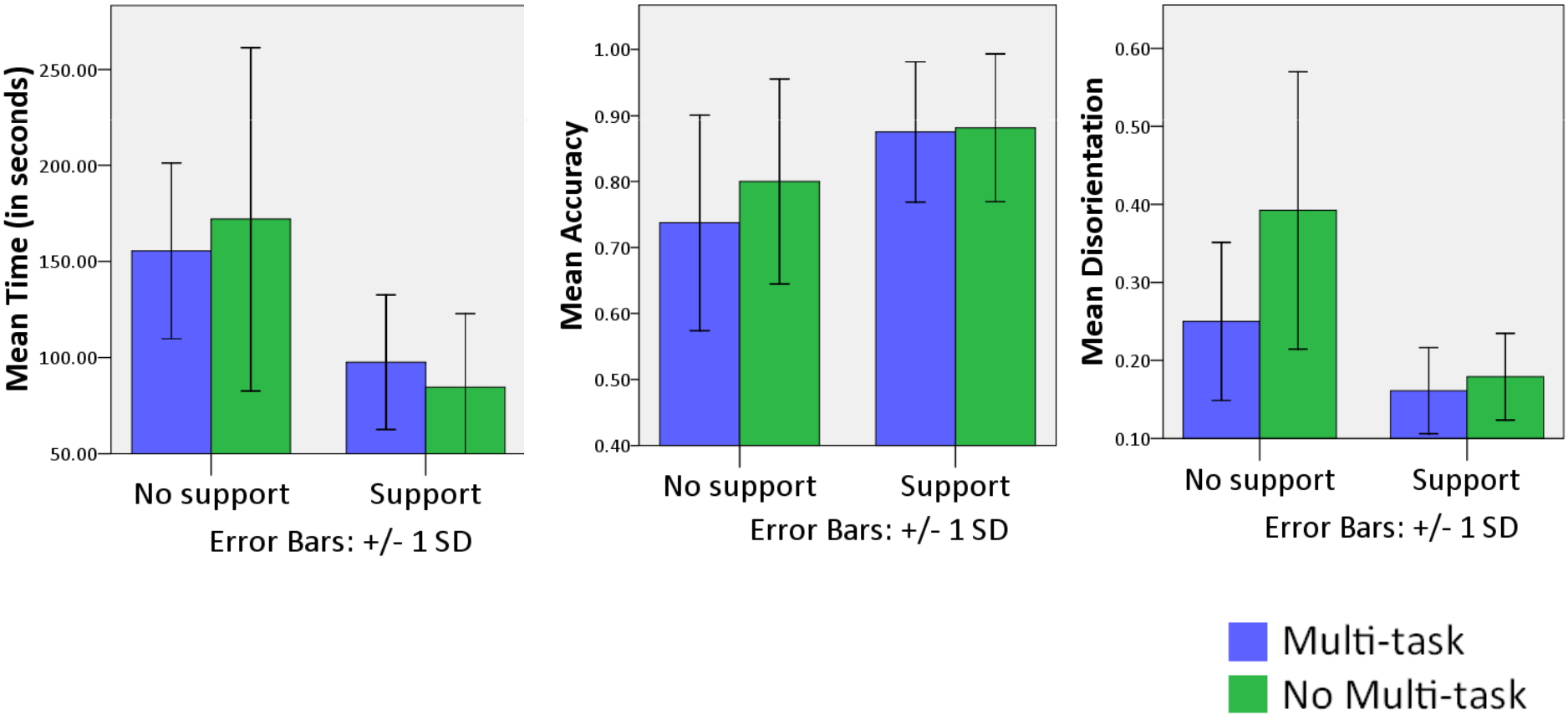
With no suggested links



With **highlighted suggested** links

Validation Experiment Results

Effect of support in Multi-task & no Multi-task conditions



Validation Experiment Results (stats)

- Three Mixed design ANOVA's were performed to statistically analyze the effects on dependent variables:

1. A strong positive effect of tool-support, F (1, 36) = 16.54, p<.001, while there was no significant effect of multi-tasking on **total time** *nor* interaction

2. A significant difference in **accuracy** between support & no-support conditions, F (1, 36) = 6.40, p<.05, while no effect of multi-tasking & its interaction with tool-support

3. Effect of support on **disorientation** was highly significant, F (1, 36) = 18.99, p <.001);
   The effect of multi-tasking was also significant F (1, 36) = 5.35, p<.05, however the disorientation was unexpectedly higher in the no-multi-tasking condition

Conclusion & Future scope

Significant positive impact of the tool-support is observed in terms of time needed to perform search tasks, disorientation in navigation and task-accuracy.

Thus, making use of cognitive model for navigation support is a useful and promising research area.

Further: Automation of the semantic feature extraction from pictures, Quantitative testing of the tool with real time websites, more experimentation with complex mock-up websites and more memory taxing secondary task.

Thank you