Goal Driven Interaction (GDI) vs. Direct Manipulation (MD), an empirical comparison

Antonio Luis Carrillo
Universidad de Málaga
Campus de Excelencia Internacional
Andalucía Tech
Department of Computer Sciences
Teatinos, 29071 Málaga, Spain
alcarrillo@uma.es

Juan Falgueras
Universidad de Málaga
Campus de Excelencia Internacional
Andalucía Tech
Department of Computer Sciences
Teatinos, 29071 Málaga, Spain
juanfc@uma.es

ABSTRACT

This paper presents a work in process about *Goal Driven Interaction* (GDI), a style of interaction intended for inexperienced, infrequent and occasional users, whose main priorities are to use a system and achieve their goals without cost in terms of time or effort. GDI basic philosophy is to guide the user about the "what" to do and the "how" to do it in each moment of the interaction process, without requiring from the user a previous knowledge to use the interface. This interaction style was introduced in previous work, where a description of its characteristics and the most appropriate user interface for it, were described. Those works included a methodology for the analysis and synthesis of the whole interactive process through a language of specification. This paper presents partial results we are collecting in real user testing, with the main aim of comparing GDI with direct manipulation interfaces (MD), nevertheless the most extended and commonly regarded as the most suitable for novice and experienced users.

Categories and Subject Descriptors

D.2.2 [**Design Tools and Techniques**]: User interfaces;
H.5.2 [**User Interfaces**]

General Terms

Human Factors, Experimentation, Design.

Keywords

User testing, Empirical usability evaluation, Guided interaction, Interaction styles, User interfaces

1. INTRODUCCIÓN

Although *Direct Manipulation (DM) with WIMP (Windows, Icons, Menus and Pointer) elements* [1] is currently the most extended desktop user interface paradigm in use, there are still many users that need training and learning period, manuals and/or expert support to become efficient users. For this reason, in previous works [2, 3, 4], the authors proposed and presented a new and alternative style of interaction: *Goal Driven Interaction* (GDI). GDI was meant to become the interaction style of choice for applications where the main priority is ease of use and minimal learning time for a user to interact with the program (as occasional, infrequent or inexperienced users), even if sacrificing speed in task achievement, the ability of running parallel tasks, and other advantages of *WIMP interfaces*.

Interacción'15, September 07-09, 2015, Vilanova i la Geltrú, Spain
ACM 978-1-4503-3463-1/15/09.
http://dx.doi.org/10.1145/2829875.2829892

The fundamentals of GDI trace back to the works of Newell and Simon [5] that were devoted to the mechanism of human reasoning for problems resolution. Their vision of problem solving (as in GDI) was based in the breaking up of the main or general goal in a hierarchical tree of sub-goals, whose branches would have different lengths depending on the degree of their fragmentation into sub-goals. The leaves of the tree would be elementary actions or sub-goals.

Based on these works, Card, Moran and Newell [6, 7] developed the most important of the existing cognitive models, the Human Processing Model, whose initial paradigm (as in GDI) consisted in conceiving the interaction as a problem resolution task, and described a psychological model of humans formed by three interactive systems: perceptive, motor and cognitive, each one would have their own memory and their own processor.

This vision of the user as an information processing system, allows for the formalization of all the activities (both physical and mental) that take part in that task, and gave origin to the methods for modelling, specification and evaluation of the user interface that are widespread today, the GOMS (Goals, Operators, Methods, and Selection rules) methods [8, 9]. Among other things, that methods allow the description of the sequences of behavior and knowledge that the user need to correctly interact with the system and accomplish his goals. The models themselves are framed in the set of techniques that allow for a hierarchical task analysis, as their main goal is the decomposition of those tasks, so that the resolution method can be followed step by step. Authors extended the NGOMSL notation [10] to use it as a source specification language, which after a compilation process, generates the corresponding GDI user interface.

Therefore, because of this extended NGOMSL models include the hierarchical knowledge the user must have (the tasks to do, and the procedures to be followed), the aim of GDI is to preclude the user from having to devote time to acquiring such knowledge. GDI's main strategy is to let the interface gradually provide the user with such knowledge, and to guide him in a hierarchical and progressive way through the whole interaction process, not only as far as the tasks and goals approach are concerned, but also about the sequence of steps to follow or the choices that can be made at any moment to achieve those goals. This strategy is, in many respects, the opposite of the MD in the sense that the user has no freedom of interpretation or any possibility of experimentation with a metaphor. As *wizards* oriented interfaces, GDI gives the control to the system, as opposed to MD that gives the control to the user.

The user interfaces based on this interaction style will need an area (as seen in **Fig. 1**) called *Goals Driver Window* (GDW) that will be the place where the user will be presented with either the *steps* of the *method* to be followed or the different *alternatives* to *choose* from, to satisfy the specific *goal* at any moment. Then, GDW becomes the substitute (or alternative) to the typical menus, icons, toolbars and those elements in WIMP interfaces, that are not necessary in GDI (as seen in Figure 1).

2. TEST CONDUCTING

2.1 Introduction

To perform the comparison between the two interaction techniques (GDI vs MD-WIMP) a set of tests involving real users is being conducted [12, 13]. For these tests a familiar task was chosen, a task for which is likely the user has some experience or at least will understand it quickly. In particular, the theme chosen for this testing is the design and furnishing of kitchens, both because it is an infrequent scenario, except for professionals in this field; and, as we have said, is a familiar task domain for almost all users, at the same time requiring manipulation of diverse objects, reasons which, in principle, could be seen favouring MD interaction style.

Two prototypes in Java language, one with the GDI interface **(Fig.1)** and the other with a classic MD WIMP interface **(Fig.2)**, both sharing most of the code except, of course, the sections in which user interface and style of interaction are involved. Although simplified systems, both offer the same functionality and allow the user perform the same tasks. The application is inspired by the desktop version of the similar one offered by the company IKEA [11].

Before starting, all participants sign an Informed Consent, and are meanwhile advised that the tests are taken voluntarily, free to leave any time without any justification. This Informed Consent also clarifies that these tests are not intended to make any personal or psychological assessment, but for the evaluation of the computer user interfaces involved.

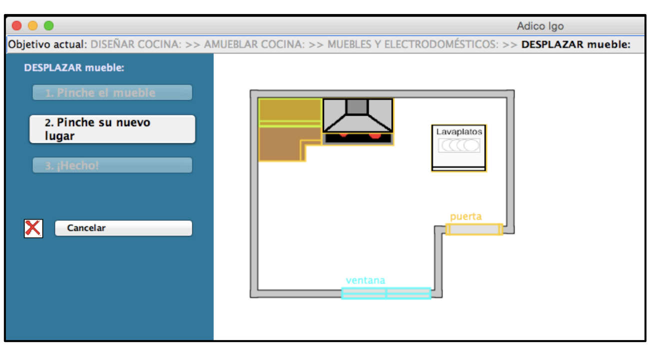

Figure 1. GDI interface screenshot

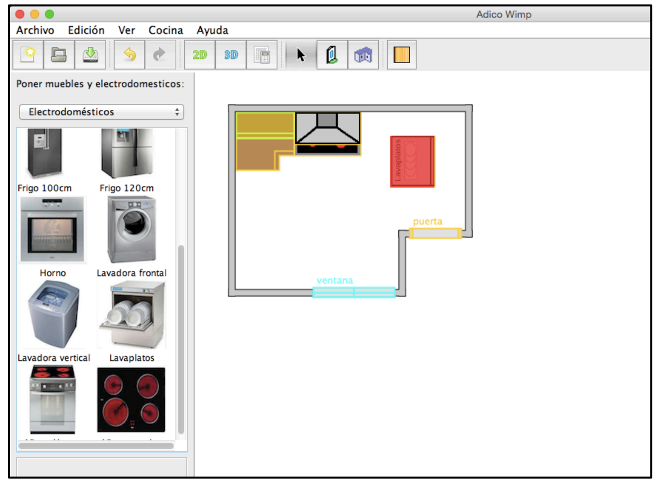

Figure 2. DM-WIMP interface screenshot

Each user should try to achieve the same goal with each of the two interfaces (GDI vs. MD-WIMP).

Data are collected both quantitative and qualitative, the former as time lengths, number of issues of assistance, number of errors (categorised by severity), etc., as well as qualitative data from the subsequent questionnaires.

2.2 User features

By the end of the study we will have tested around 25 participants. This number will include EyeTracker analysis for some of them. If any conclusions would deem it appropriate, a second phase with more users would be conducted.

All users, except a small percentage of them, will be inexperienced users. The rest will be users with an intermediate computer level, and a very small percentage of them will be kitchen design professionals, used to using specialised applications for this purpose.

The age range is planned to span from 10 to 65 years (with an average of 45 years) without any disabilities described.

2.3 Test conditions

The tests are performed individually, in controlled spots, without distractions, thus allowing both techniques Thinking Aloud as well as screen and sound recording. The user does not feel any other influence but the presence of the moderator and the applications under test.

The estimated time per user (including questionnaires) ranges between 45 to 75 minutes.

The whole process of interaction is collected with two recordings. The first one is taken with a external video camera, located next to the user recording the user hands and the screen, and any incidences, assistances, etc., for further re-examination, if necessary. And the other recording, even more useful, is that which is taken with computer screen recording software.

A common laptop computer is used for all tests, which facilitate their relocation and is less intimidating and more familiar to users than any other complex installation. External video recording is also more conveniently done with a small camera.

2.4 Test development

The tests follow the next steps:

1. The user fills out a preliminary questionnaire containing personal information such as their age, genre, experience with technological and digital devices, domain knowledge of the task, etc.
2. The user receives a single sheet briefly describing the three group of tasks to do **(Fig. 3)**. This sheet with the tasks is available to the user throughout all the process.
3. Then, following a *balanced* strategy *within-subjects*, each user is made to use both versions of the application, alternating the order among users. *Thinking Aloud* technique is used in a relaxed way, not making at any time the user to explain what is being doing, but leaving them freedom to give explanations when necessary. The moderator tries not to interfere at any moment, or if so, only in critical or blocking situations. In any case the incidents are reflected and graded as part of the data collected.
4. At the end of each of the two tests (each one corresponding to a type of interface), and before moving on to another type of interface, the user fills out a questionnaire with 10 questions on specific points of the process that has just experienced. This covers important aspects for the final evaluation.
5. Finally, after the two tests and the corresponding questionnaires are done, the user fills out a quick questionnaire comparing both styles of interaction. Among the questions asked are: with which interface have you had less hesitation about *what* the next step was; with which was it easier to know *how* to take every step; for which do you consider that more aided should be included; which one seemed easier to use and required less training; with which was it faster to operate; and, naturally, their final preference.

3. EVALUACIÓN DE RESULTADOS HASTA EL MOMENTO

The **Table 1** displays the partial empirical results of the first 8 users. The last columns of the **Table 1** display the types of incidences and, perhaps, the need of assistance by the moderator. There are three types of incidences: 1) slight, 2) non-locking, and 3) severe or locking. The most severe is the third type of incidence, which takes place when the user is stuck with something in the interface. This would prevent the user from finishing the task. This will require the assistance of the moderator in order to exploit the collected data and analyse the rest of the planned tasks for that user.

Non-locking issues correspond to detected or undetected mistakes that somehow affect the final results. For example a common error is repeatedly confusing the 'undo' action with the deletion of the last object added; or forgetting performing a task, or initially not choosing the correct geometry, erroneously keeping throughout the process the wrong shape of the kitchen, and perhaps requesting a warning by the moderator.

Finally, slight incidents could not be properly considered errors because could be sorted out by the user, perhaps with some delay. They are non-locking and did not need any moderator assistance.

An initial analysis of the empirical data obtained until this document was sent may be premature, nevertheless, two apparent tendencies can be drawn from those data: 1) GDI is more efficient in time (took less time in accomplishing the whole tasks), and 2) GDI needed fewer moderator warnings or assistances. On the other hand, from the questionnaires is explicit that all users have a sharp preference for the GDI interface. This might not be surprising among inexperienced users, but it was not expected for users of an intermediate computer level, and especially surprising in the case of professionals on kitchen design, who are used to more complete and complex MD interfaces. In addition, all users in the final questionnaire, that includes to compare both styles of interfaces, clearly opt for the GDI guided interface.

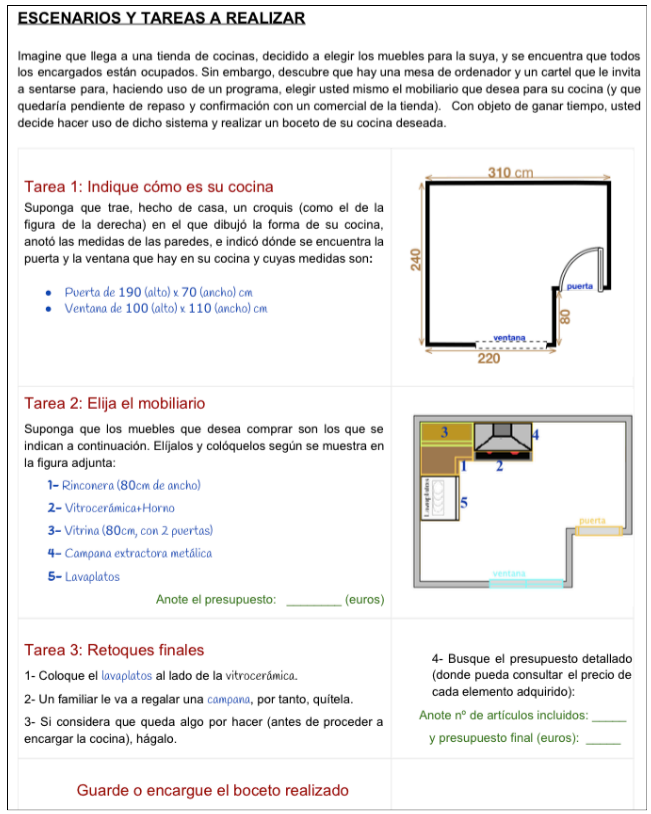

Figure 3. Sheet with the tasks handed out to users

4. CONCLUSIONS

The work in process presented here complements the theoretical study of the GDI. This style of interaction is intended for applications that require the minimization of the learning time up to a point that makes them appropriate for inexperienced, infrequent or occasional users. The essential point here is no requiring previous knowledge. With this aim, these interfaces should guide the user towards their goal. To do so, indications of both the "what to do" and the "how to do it" must always drive the user at all times.

The partial results obtained so far show that, compared with a MD-WIMP interface, the total execution time is shorter, there is a fewer number of incidences, and finally the users show a clear preference for this guided interaction style.

Table 1. Partial empirical results (time and incidences) of the first 8 users.

User / (category)	Evaluation order	Task 1 Time		Task 2 Time		Task 3 Time		Total Time		Incidences		
										Slight	Non-locking	Locking
USER 7 (inexperienced)	**GDI**	6:20		9:20		8:11		23:51	✔		1	
	WIMP	14:23		11:26		11:32		37:21 (∞)			1	2
USER 2 (inexperienced)	WIMP	7:05		11:30		7:00	∞	25:35 (∞)			3	3
	GDI	5:30		7:00		3:30		16:02	✔			
USER 6 (inexperienced)	**GDI**	6:11		8:07		6:17		20:35	✔			
	WIMP	4:40		5:18		17:55	∞	27:53			3	
USER 4 (inexperienced)	WIMP	9:50		13:20		12:50	∞	36:00 (∞)	∞	2	2	3
	GDI	8:30		4:40		4:46		17:56	✔			
USER 1 (intermediate)	**GDI**	2:45		3:35		2:30		8:50	✔	1		
	WIMP	2:30		2:36		4:14		9:20			3	
USER 5 (intermediate)	WIMP	4:10		4:05		6:45		15:01		1	2	
	GDI	2:53		4:07		4:10		11:10	✔		1	
USER 3 (kitchen design professional)	**GDI**	2:30		4:15		1:45		8:30	✔			
	WIMP	5:55		1:50		2:00		9:45		1	2	
USER 8 (kitchen design professional)	WIMP	4:34		3:40		3:53		12:07				
	GDI	2:18		2:16		3:23		7:57	✔			

5. REFERENCES

[1] Dix A., Finlay J., Abowd G., Beale R., 1998: *Human-Computer Interaction*. 2nd ed. Prentice-Hall.

[2] Carrillo A., Falgueras J., Guevara A. 2005. A notation for Goal Driven Interfaces specification. In *Interacción 2004*, Raquel Navarro-Prieto & Jesús Lores Eds. Springer, Dordrecht, The Netherlands.

[3] Carrillo, A., Guevara. A., Gálvez. S., Caro. J. 2002. Interacción guiada por objetivos. In *Proceedings of the 3th congreso internacional Interacción Persona Ordenador* (Leganés, Spain, May 8-10, 2002). INTERACCION'02. 68-75. ISBN: 84-607-4501-5

[4] Carrillo A., Falgueras J., Guevara A. 2006. GDIT: Tool for the design, specification and generation of Goals Driven User Interfaces. In *Proceedings of the 8th International Conference on Enterprise Information Systems* (Paphos, Cyprus, May 23-27, 2006). ICEIS'06.

[5] Newell, A., Simon, H. 1972. *Human Problem Solving*. Prentice-Hall.

[6] Card, S., Moran, T., Newell, A. 1980 Computer text-editing: An information processing analysis for a routine cognitive skill. *Cognitive Psychology*, 12, 32-74.

[7] Card, S., Moran, T., Newell, A. 1983. *The Psychology of Human-Computer Interaction*. Hillsdale, New Jersey: Erlbaum.

[8] John, B.E., Kieras, D.E. 1996. The GOMS family of user interface analysis techniques: Comparison and contrast. *ACM Transactions on Computer-Human Interaction,* 3, No.4, (December 1996), 320-351.

[9] John, B.E., Kieras, D.E. 1996. Using GOMS for User Interface Design and Analysis: Wich Technique? *ACM Transactions on Computer-Human Interaction*, 3, No.4, (December 1996), 287-319.

[10] Kieras, D. 1997. A Guide to GOMS Model Usability. Evaluation using NGOMSL. In *The handbook of human-computer interaction*. (Second Edition), M. Helander & T. Landauer Eds. Amsterdam, North-Holland, 733-766.

[11] Ikea, 2015. Last access: March, 2015: http://kitchenplanner.ikea.com/ES/UI/Pages/VPUI.htm

[12] Barnum, C. M. 2011. Usability-Testing-Essentials. Elsevier

[13] Dumas, J. S., Redish, J.C. 1999. *A Practical Guide to Usability Testing*. Intellect Book.