

1991.10.18.

정보소통 공간과 투명한 코드, 투명한 컴퓨터

金 興 圭 (고려대)

1. '투명함'의 의미

- 1) 기본적 의미: 물체 X가 인식 주체(A)와 대상물(B) 사이에 있으면서 그 인식 활동을 차단하거나 방해하지 않을 때, X는 A→B 사이에서 투명하다.
- 2) 확장된 의미 1: 주체(A)가 어떤 도구(X)를 써서 일정한 대상(B)을 다루는데, 도구의 기능이 원활하여 주체 자신의 연장(延長)처럼 자명할 때, X는 A→B 사이에서 투명하다.(도구의 투명성)
- 3) 확장된 의미 2: 주체(A)가 어떤 중간과정(X)을 거쳐 목적 단계(B)에 이르는 데, 그 중간과정이 독립적 활동으로서의 가치를 지니지 않으며 특별한 주의를 요하거나 지체되지 않을 때, X는 A→B 사이에서 투명하다.(과정의 투명성)

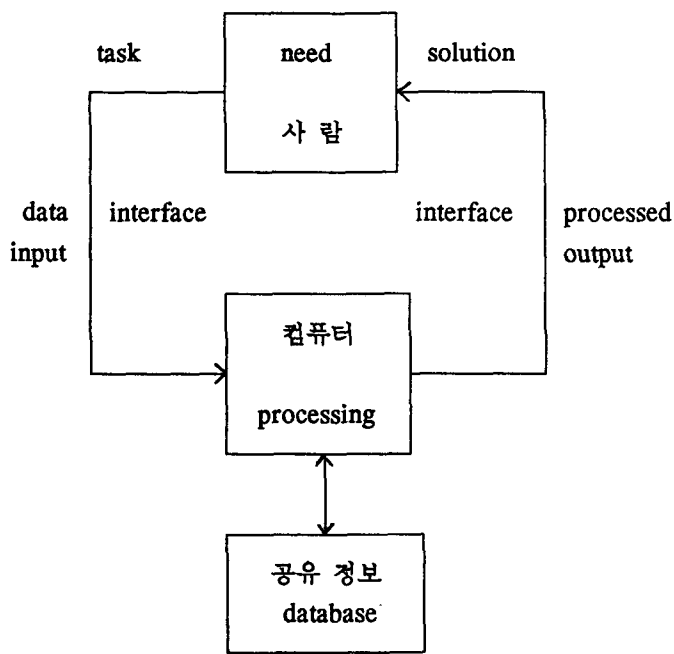
2. 오늘날의 컴퓨터, 그 불투명성

위의 2), 3)의 의미 수준에서 본다면 오늘날의 컴퓨터는 대다수의 사람들에게 극히 불투명한 도구이며, 컴퓨터 작업 역시 매우 불투명한 과정이라 말할 수 있다. 이러한 불투명성을 줄이기 위해 많은 노력이 행해지는 중이지만, 그 진전은 기대 수준보다 더딘 것으로 보인다. 가장 진보된 지식을 산출하며 풍부한 지적 자산을 조직·관리한다는 대학에서조차도 대다수의 대학인들에게 컴퓨터는 불투명한 물건이며, LAN은 불투명한 迷路처럼 여겨진다.

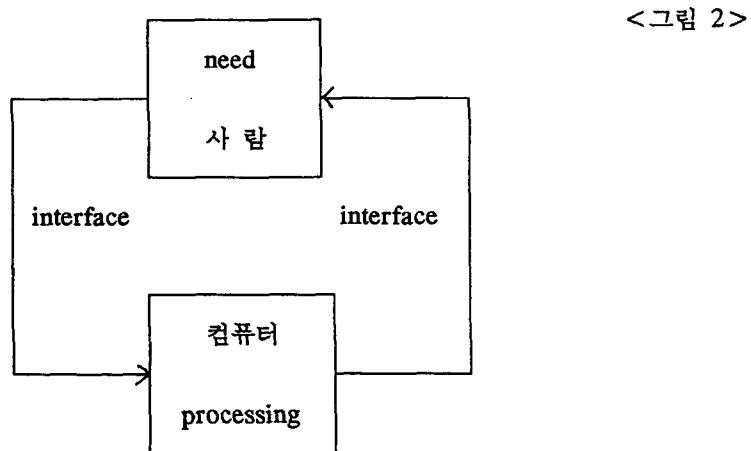
거두절미하고 말한다면, 이러한 불투명성의 해소가 더딘 이유의 중요한 일부분은 컴퓨터와 컴퓨팅에 관한 접근 패러다임의 편향에 있다고 생각된다.

'정보 처리 장치'로서의 컴퓨터 시스템과 사람 사이의 관계를 다음과 같은 그

림으로 추상화해 볼 수 있다.



오늘날의 전산학, 정보처리 연구 및 컴퓨터 산업은 위의 그림 가운데서 큰 사 각형의 안쪽에 있는 사항들에 편향된 추세를 보이고 있다. 그것을 그림으로 나타 내면 다음과 같다.



PC 사용이 대중화되면서(다시 말해 PC 시장 규모가 엄청나게 커지면서) 사용

자가 쉽게 친해질 수 있는 컴퓨터와 user interface가 중요한 관심사가 되었다. 여기에는 물론 사용자의 요구가 작용했지만 그보다 직접적인 요인은 컴퓨터 생산-공급자쪽의 마케팅 동기가 아니었던가 한다. 컴퓨터에 낯설거나, 불친절한 기계와 어려운 OS에 질린 고객들의 지갑을 열게 하기 위해서는 user friendly interface가 가장 좋은 세일즈맨이었기 때문이다. 이러한 추세는 연구 재정의 조성과 배분이라는 메카니즘을 통해 컴퓨터 및 인접 분야의 연구에도 크게 작용하고 있는 것 같다.

이렇게 말한다고 해서 user interface의 의의를 부정하거나 적게 평가하려는 것은 아니다. 더 친근하고, 더 자연스러우며 편리한 컴퓨팅 환경은 '도구로서의 컴퓨터의 투명성'을 높이는 데 아주 중요한 요소다. 우리가 문제로 여기는 것은 그것을 향상시키기 위한 노력 자체가 아니라 그것에의 '편향'과 이에 따라 나타나는 바 '연구 인력, 비용 및 사회적 관심의 배분에 있어서의 불균형'이다. 다시 말해서 우리는 'how to interface'에 치중하는 동안 'interface for what'라는 질문에 소홀하지 않았나 하는 것이다.

3. 정보 중심적 접근

위에 말한 투명성 저편에 있는 대상물 내지 목표는 물론 정보이다. 컴퓨터의 투명성을 증진하기 위해서는 사용의 편의성 못지않게 목표의 가치를 높이는 일이 중요하다. 목표 가치가 미미하거나 흐릿할수록 그 사이에 있는 물체의 불투명성은 커지기 마련이다.

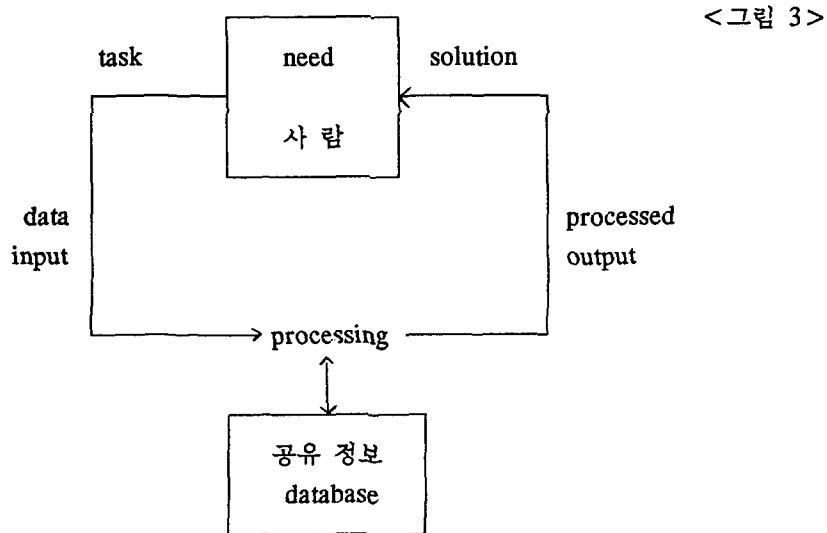
누구나 아다시피 컴퓨터 하드웨어와 소프트웨어의 비용은 전자가 대부분을 차지하던 단계에서 후자가 압도적인 비율을 차지하는 단계로 이미 변화했다. 앞으로의 세계에서는 여기에 정보(반가공 혹은 완전가공되어 데이터베이스로 활용되는 정보)의 비용이 중요한 요소로 들어오고, 마침내는 가장 큰 비중을 점하리라는 관측이 유력하다. 특히 미래의 컴퓨터는 텍스트, 음성, 圖面, 음악, 영상 등의 정보를 처리하고 그러한 데이터베이스의 그물 속에서 움직이며 통신하게 될 터이므로, 잘 집적되고 효율적으로 관리되는 정보체계는 지금 상상하는 정도를 능가하는 중요성과 사회적 자산 가치를 띠게 될 것이다.

이러한 추세에 직면한 현재의 시점에서 우리나라의 정보산업은 어떠한가? 최근의 <DB산업 육성을 위한 세미나>에 나온 자료에 의하면 한국의 데이터베이스 수는 100여개로 일본의 1/25, 미국의 1/35이며, 매출액은 각각 1/27, 1/33 정도라고 한다. 지난 9월 초 동경에서 'DATABASE '91 TOKYO' 라는 데이터베이스 전시회가 열렸는데, 이에 관한 보도를 보면 일본의 데이터베이스 종류는 총 2,100에 1년간 매출액이 1,600억엔이다.

정보의 수집.구성.전달.보존이라는 측면에서 인류의 문명사를 살펴 보면 몇 단계의 중요한 도약이 있었다. 문자의 발생, 인쇄술의 발명, 有.無線 통신과 전파 매체의 등장 등이 그것이다. 컴퓨터 네트워크에 의해 매개되는 다중 매체(Multi

Media) 정보 시대는 이들 모두를 흡수하면서 지식.정보.자료의 축적.보관.검색.분배에 있어서 또 한 단계의 획기적 변화를 가져올 것이다. 이러한 예상은 사실 상식적인 것에 불과하다. 신문의 과학면 정도를 읽는 사람이라면 누구나 이 예측에 동의한다. 그러나 예상하는 사람은 많아도 그것을 위해 준비하는 노력은 적은 것 같다. 컴퓨터 및 정보처리를 다루는 우리의 관심은 아직 '사람과 기계의 매개'(Man-Machine interface)에 주력하는 도구 중심적 접근(tool oriented approach)이 우세하고, 정보의 축적.구성.검색.분배 및 이에 수반된 기술적, 심리학적, 철학적, 문화적, 사회적, 경제적 문제를 연구하는 정보 중심적 접근(information oriented approach)은 빈약하다고 생각한다.

이를 극복하기 위해서는 앞에서 제시했던 <그림 1>의 모형을 <그림 2>처럼 도구 중심적으로 단순화할 것이 아니라, 오히려 다음의 <그림 3>처럼 정보 중심적인 시각에서 보도록 하는 일이 필요하리라 생각한다.



4. 투명한 코드

컴퓨터와 마찬가지로 코드 또한 투명해야 한다. 정확하게 말하자면 코드의 투명함이 없이는 컴퓨터가 투명해질 수 없다.

필자는 6년 전까지 타자기로 글을 썼다. 타자에 숙달되면 타자기도 앞에 말한 의미에서 투명해진다. 글을 쓰는 동안의 의식 속에 타자기는 존재하지 않고 '글과 나'만이 있다. 그러나 "넋"이라든가 "스름의" 같은 글자가 나오는 순간 타자기는 거북하고 불투명한 물체가 된다. KSC-5601라는 '표준코드'를 갖춘 오늘날의 컴퓨

터는 한국어에 대해서 그것보다 더 고집 세고 불투명한 물체 노릇을 한다.

2바이트 조합형이나 완성형이나 하는 논의를 여기에서 되풀이하자는 의도는 없다. 2바이트 조합형 한글 코드로도 옛글자까지 완벽하게 처리할 수는 없고, 한자 처리의 제약이 큰 점은 여전히 속계로 남는다. 더우기 한글(옛글자 포함), 한자, 영어, 러시아어, 불어, 아랍어 등이 섞여 있는 데이터베이스를 만들어 자유로이 입출력하기는 불가능하다. 하기는 우리만 그런 것이 아니라 미국인들도 현재의 8비트 코드 체계로는 印歐語, 슬라브어, 아랍어권의 각종 텍스트 모두를 자연스럽게 수용하는 수가 없다. 'escape sequence' 등을 사용하는 편법은 어디까지나 편법이며, 정보 교환의 투명성이 보장되지 않는다.

그렇다면 왜 영어·독어·불어 문자들은 1바이트여야 하며, 한글·한자·가나는 2바이트여야 하는가라는 질문이 필요하다. 이들은 각각 2바이트와 4바이트로 되어도 무방하다. 지금까지 그렇게 하지 못한 이유는 다음의 두 가지이다.

<1> 그렇게 하자면 컴퓨터 메모리와 저장 장치에 두배의 비용이 든다.

<2> 지금까지의 정보 처리는 한 나라 혹은 하나의 문자권 안에서 이루어졌고, 그러한 장벽을 넘어선 다중언어 정보의 축적과 공유가 절실하지 않았다.

이 가운데서 <1>의 요인은 이제 대단치 않은 것이 되었고 앞으로는 더욱 그러할 것이다. <2>의 요인은 지식 데이터베이스와 정보 소통의 국제적 통합이라는 조류로 인해 더 이상 지속될 수 없다.

국제표준기구(ISO) 역시 이러한 필요성에 주목하면서 ISO 2022 대신 4바이트 코드 개념에 바탕을 둔 멀티 옥테트 코드(ISO DP 10646) 쪽으로 국제적 문자 코드의 통합을 추구하는 것으로 보인다. 순탄하게 될지는 모르겠으나 이 기구의 계획대로 1992년까지 새로운 통합코드 체계가 이루어진다면 ISO 2022의 아들 격인 KSC-5601은 어디로 갈 것인가? 오늘의 컴퓨터가 제대로 처리하지 못하는 우리의 문화적·지적 유산은 어떻게 처리될 것인가? 이 물음에 대한 대답도 역시 투명하지 못한 듯하다.

기술적인 부분은 필자가 말할 만한 사항이 아니지만, 최근에 몇몇 사람들에게 의해서 한글(옛글자 포함), 한자와 여러 종류의 외국 문자를 4바이트 및 2바이트 체계로 처리하는 방법이 시도되고 있는 것은 코드의 투명성 획득이라는 점에서 주목할 일이라고 생각한다. 그러한 모색들이 한정된 연구 팀이나 개인들의 분산된 노력에 그치지 않고 더 나은 공동적 해결 방안의 수립에 기여할 수 있도록 좀더 적극적인 논의가 이루어져야 할 것이다. ■