USDA Agricultural Research Service

NAME THAT FLY

Computers help make species identification child's play

By ELIZABETH PENNISI

After struggling through fish biology, counting dorsal spines or fin rays to try to distinguish daces, darters, and minnows, I would never have believed that identifying species could be fun.

But the students at Lincoln Middle School in Pullman, Wash., actually enjoy figuring out the names of their local flora and fauna. Last fall, botanist Richard Old visited that school. He first asked the students to fill out a questionnaire that included their name, birthday, favorite color, street address, and so forth. Then, by asking one student just the day of the month in which he was born and his preferred color, Old was able to home in on that student's name out of a pool of 180.

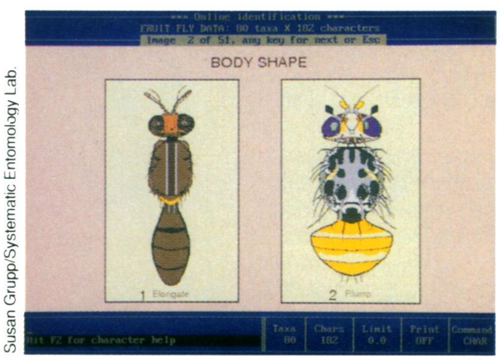
The students were dumbfounded. They couldn't wait for Old to do it again. For them, "it was like watching card tricks," Old recalls. "Then I told them they could identify plants with the same power. They were hooked."

He proceeded to help the kids build an identification system — called a "key" — for all the plants in their school yard, letting them pick the "traits" and afterward letting them identify, or "key out," specimens based on these traits. Typically, keys make use of many technical terms. But for the students, "smells bad" or "feels squishy" worked as well as "pinnate foliage" or "clustered heads" as distinguishing characteristics — thanks to a computer program created by Old. The program creates identification keys for all sorts of purposes and for use by all sorts of people. With this system, one can leave out the technical descriptions that stymie all but a few experts.

Like me and many other biology students, Old remembers all too well the frustration of keying out plants or animals he found. Once, he collected a 15-foot-tall grass specimen. Although he knew that the plant's height made it unique, the identification guide forced him to examine the grass under the microscope before it led him to its name. "It just made me mad," recalls Old, now a botanist who does plant identification at Washington State University in Pullman.

Later, while working as an agricultural extension agent, he often found himself predicting a weed's identity based solely on a farmer's coarse description of the weed's height, habitat, and flower color. He knew, for example that a 6-foot, purple-flowered plant that grew near water was purple loosestrife.

"That description is not good enough to identify any plant anywhere in the world, but I knew what [the farmer] was talking about. So I decided to put that ability into a system," Old explains. He designed an expert system, a computer program that works the way he thinks. Old later created a company, XID Services, Inc., in Pullman, to market this technology.



Susan Grupp/Systematic Entomology Lab.

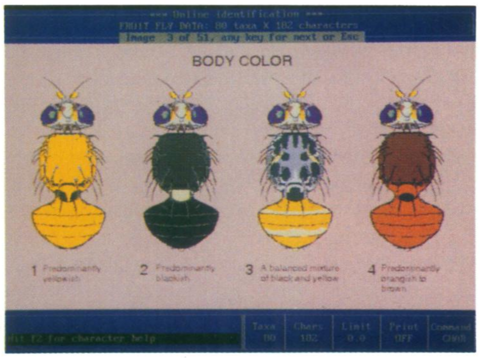
In expert systems, pictures make identifying the Oriental fruit fly a much easier task.

This winter, federal biologists joined Old in the push to computerize two jargon- and data-laden scientific disciplines, taxonomy and systematics, the studies of the classifications of organisms. In one case, the U.S. Department of Agriculture (USDA) has readied an expert system, complete with pictures, for identifying fruit flies. Elsewhere, both individuals and institutions are reworking species lists for on-line use (see sidebar).

This electronic revolution will make these disciplines and the knowledge they yield more accessible. Electronic communication can get data out faster than printed scientific journals, and expert systems can present those data in ways nonexperts can understand.

The emphasis on computerization also benefits those who spend long hours examining specimens in the hope of naming and determining the relationships among species, says F. Christian Thompson, an entomologist at the USDA Systematic Entomology Laboratory at the Smithsonian Institution in Washington, D.C. Lacking the glamour and income-generating potential of research fields such as molecular biology, systematics has attracted less interest. Consequently, there are fewer opportunities for students to train for this work, Thompson says.

Yet the need for this expertise is expanding. Each year, Thompson and his colleagues identify more than 100,000 "unknowns" sent to them by state and federal agencies, universities, and companies both in the United States and abroad. In other laboratories throughout the world, thousands more "unknowns"



Grupp

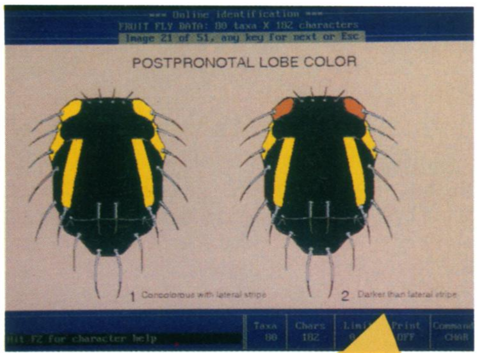
are piling up on systematists' workbenches, creating a backlog of work that slows the assessment of biological diversity. Thompson is convinced that the development of easy-to-use computer programs can reduce some of this backlog by enabling people who are less well trained in taxonomic principles to identify more organisms on their own.

"I do believe it's going to be the way things are identified in the future," Old adds.

Old programmed his expert system to follow the logic people apply to narrow their choices of an organism's identity. With each trait, or character, considered, the computer rules out all but those plants (or animals) with that trait as well as all the preceding traits already registered.

"It's a much more efficient way of getting to the classification," Old says. It's also quite different from dichotomous keys, traditionally used to identify species. "Dichotomous keys are very rigid and very cumbersome," says Old. Rather than use a key correctly, "most people just flip through the pages," he adds.

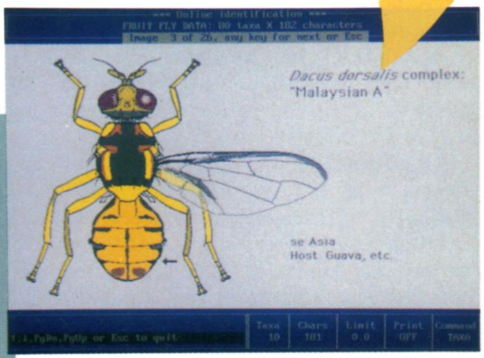
Unlike written keys, which work only if traits are considered in a specified order, expert-system keys use whatever characteristics one



Grupp

knows — in whatever order one presents them — to home in on a plant's name. Because of the way his software is set up, Old says he can enter up to 500 traits about a new plant in less than 20 minutes. The setup lets him squeeze a data set with 500 characteristics for each of 1,000 plants into relatively little space in the computer's memory.

The program also provides helpful hints. If the person describing the plant runs out of ideas about what characters to consider, the computer evaluates the



Britt Griswold/Systematic Entomology Lab.

possible choices, then figures out and asks about a new trait that will easily distinguish one of those choices. Should someone not understand, for example, the difference between a spine and a prickle, the computer's "help" file will describe them in words nonscientists can understand. A pointed stub, rigid enough to break the skin and draw blood, qualifies as a spine, while softer stubs are prickles, it says.

Already, several other researchers have seen the potential of this approach. Field scientists collecting tropical insects hope that having portable computers at their collection sites will speed classification. Another researcher wants to develop such a system for identifying poisonous plants, Old says.

Thompson and his colleagues have made expert-system keys even more user-friendly. Created for

Continued on p.111

Cataloging the world's biota — electronically

As of January, it takes just a few keystrokes to tap into the 88,000 most commonly requested plants in the type-specimen collection of the U.S. National Herbarium. For animal lovers, there's the list of the 4,629 currently recognized mammals. One can catch up on conservation efforts by pulling up onto the computer screen the BIOLOGICAL CONSERVATION NEWSLETTER AND BIBLIOGRAPHY. For more esoteric reading, try the newsletter of the American Society of Plant Taxonomists.

These resources, as well as a checklist of plants in the Guianas, a guide to the National Herbarium's historical collections, and a computer program that helps researchers assess the relatedness of species, have become available through the Smithsonian Institution's Natural History Gopher Server.

Gopher servers are rest stops along that booming electronic highway called Internet. The gopher software enables users to gain access to databases at different institutions. By developing electronic databases, the National Museum of Natural History's Collections and Research Information System Program expects to keep the information more current and to make it more readily available to scientists and the public.

The gopher system is just one of many efforts under way to put the world's plants and animals, and in some cases their natural histories, on-line. In late 1993, Congress appropriated money for the creation of a new Interior Department bureau, the Biological Survey, which will devote considerable resources to putting information about U.S. plants and animals into electronic databases. Already, states are gathering and computerizing these kinds of data (SN: 10/16/93, p.248).

It's an awesome task. For example, USDA's F. Christian Thompson is helping to compile a "telephone directory" of all the world's insects. His volume of the directory, which deals only with the 100,000 or so species of flies — a small percentage of the world's millions of species — will include the scientific name of the insect, the name of the person who first described it, and where and when the species was discovered. It will also give the name and phone number of an authority on that species and detail its classification. Within the next year, Thompson hopes to have complete information for North American flies.

"Then," he says, smiling, "we'll go for the world." *— E. Pennisi*

Environment

Organochlorines lace Inuit breast milk

The Inuit of northern Quebec dine on seal and beluga whale blubber — food loaded with organochlorine compounds such as the pesticide DDT and PCBs (polychlorinated biphenyls).

Probably as a result of this diet, Inuit mothers exhibit some of the world's highest recorded concentrations of PCBs in breast milk, report Pierre Ayotte and his colleagues at Laval University Hospital in Ste-Foy, Quebec. But omega-3 fatty acids in the blubber may help protect against some of the organochlorines' toxic effects, Ayotte adds.

The milk of the 107 Inuit studied had organochlorine pesticide and PCB concentrations four to seven times higher than the breast milk of Quebec women who don't eat blubber, Ayotte and his team write in the December Environmental Health Perspectives.

Though studies have shown that such high concentrations of PCBs in breast milk can impair brain development, Ayotte says his preliminary studies of Inuit children indicate that they are developing normally. He speculates that the mothers' consumption of a diet rich in omega-3 fatty acids may protect against the damage to the central nervous system caused by organochlorines.

Still, the Inuit may pay a price for eating organochlorine-rich food. High rates of infectious disease among their infants may stem from PCB-related immune-system damage, the researchers write.

DDT, certain PCBs, and other organochlorines also possess estrogenic properties. In the Feb. 2 Journal of the National Cancer Institute, another Laval University group reports on a small study showing that breasts of Quebec-area women with estrogen-responsive cancer tend to be more heavily contaminated with DDE, a breakdown product of DDT, than breasts of women with tumors unresponsive to estrogen. The findings of this team, led by Eric Dewailly, support the idea that estrogenic organochlorines may foster hormone-responsive breast cancers (SN: 7/3/93, p.10).

EPA wants close scrutiny of chlorine

In its proposal for revamping the Clean Water Act, the Environmental Protection Agency has recommended examining chlorine's impact on health and the environment (SN: 1/22/94, p.59) — with the possible goal of banning or restricting its use, EPA officials said last week.

Congress must pass an amendment this year to reauthorize the law, which otherwise would expire. Sen. Robert Graham (D-Fla.) has said he intends to introduce on Feb. 23 an amendment that closely resembles EPA's plan.

The agency's proposed $2 million, one-year chlorine study would look at the effects of the use of chlorine and chlorine compounds in the manufacture of paper, solvents, and plastics and in disinfecting waste water and drinking water, says EPA's James F. Pendergast. It would also assess the availability, effectiveness, and safety of chlorine substitutes.

Another study EPA calls for would test the environmental and economic results of pollution trading. For example, the plan says EPA might consider allowing an electric utility whose nitrogen dioxide emissions pollute a watershed to meet its emission standards by paying farmers to reduce their use of nitrogen-rich fertilizers.

The agency also proposes to study how to control the ill effects of runoff from agricultural irrigation and to analyze the costs and benefits of the Clean Water Act, Pendergast says.

Continued from p.109

the USDA Animal and Plant Health Inspection Service, their program can thus far identify 67 fruit fly species. Eventually, the system will include data on the 200 most troublesome fruit flies USDA inspectors are likely to come across.

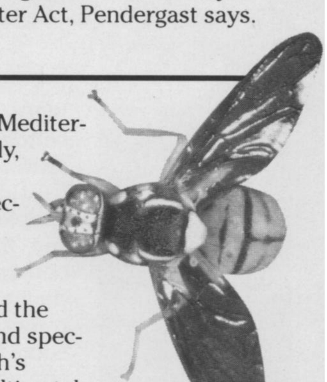
First, the group collected the necessary taxonomic information and reference data for these 200 insects. The researchers initially thought they could scan photographs or specimens with a video camera and load those images into the computer's memory. But when the project got off the ground four years ago, available technology did not provide enough resolution. So Thompson and his colleagues turned to artists' renderings. The finished program will include hundreds of drawings, of both whole insects and key features used in the identification process. Menus on the computer screen enable even a neophyte to work through the program.

With dichotomous keys, one can arrive at the insect's name through only one route, says Thompson. Thus a single mistake can lead to an erroneous identification. But the computer works backward, forward, or circuitously, narrowing the choices based on traits considered in any order.

In addition, the computer program will allow the mismatch of one or two characters in the decision-making process and then help the user "correct" the mistake without having to start all over again. Finally, the fruit fly expert system gives detailed information about the insect once it is identified and provides ways to verify the identification, Thompson adds.

With this program, Thompson hopes to reduce greatly the number of "unknowns" he receives. And that pleases plant pathologist Rebecca A. Bech. As a coordinator of USDA inspectors who patrol the nation's borders, she can't wait until this system is up and running. "We're the first-line barrier to keeping these plant pests out," Bech says.

At each location, inspectors must be able to pick out exotic mollusks, insects, even seeds that cross U.S. borders. At ports, the inspectors board ships to take a close look at the cargo and packing material. When that cargo consists of produce, "quite often they will find flies flying around," says Thompson. In 1992, of 37,467 pests intercepted, 4,625 were fruit flies. Missing these pests can be devastating: For the 1993 to 1994 budget, California allocated $8.1 million to eradicate just one, the Mediterranean fruit fly, Bech notes.



If the inspectors do not recognize the insects, they impound the cargo and send specimens to Bech's group — or ultimately to Thompson's lab — for identification. This causes costly delays in unloading, especially for highly perishable cargo. "If it's a Mediterranean fruit fly, they are going to destroy everything," says Thompson. Inspectors also must set up sampling traps near the ship's dock to monitor whether an infestation has begun.

Not too long ago, for example, a USDA inspector found fruit flies in ornamental pepper plants in a cargo of fresh-cut flowers. Because produce rarely passed through that port and because the ship came from a place that was not supposed to have these insects, the inspector impounded the cargo and sent specimens across the country for identification. It did, indeed, turn out to be a pest that required destroying the plants.

Says Bech: "If we had the fruit fly expert system, we would have been able to get on this immediately." □