

Copying As Text

INTRODUCTION

This chapter describes *Expressionist*'s Copy As Text facility, which allows you to copy *Expressionist* equations out as text (rather than pictures), for use with other applications.

Typically, typesetting systems create aesthetically pleasing equations, but lack a simple means of creating these expressions. *Expressionist* permits you to build complicated equations in a WYSIWYG manner, then use Copy As Text to move the equations to a text-based system.

Copy As Text (herein also referred to as CAT) creates a textual description of a graphical equation. This description may be used to create complex equations in traditional text-based typesetting systems (e.g., \TeX , troff, etc.) or word processors which have proprietary equation markup languages, such as Lotus Ami Pro and WordPerfect.

Expressionist can translate WYSIWYG equations into various text formats, including several varieties of \TeX , troff's eqn, and PostScript. Each format requires a translation table, or CAT file, that describes the rules of the particular markup language.

In general, copying text out of *Expressionist* is done this way:

1. Choose the desired CAT Translation Table.
2. Make *Expressionist* Copy As Text (instead of Copy As Picture).
3. Select the expression to be copied.
4. Copy the selection to the clipboard.
5. Paste into the desired document (e.g., \TeX document).

The specific steps are described in the following Quick CAT Tour.

SUPPLIED TRANSLATION TABLES

Expressionist includes many different translation tables, known as CAT files, for use with various applications and markup languages.

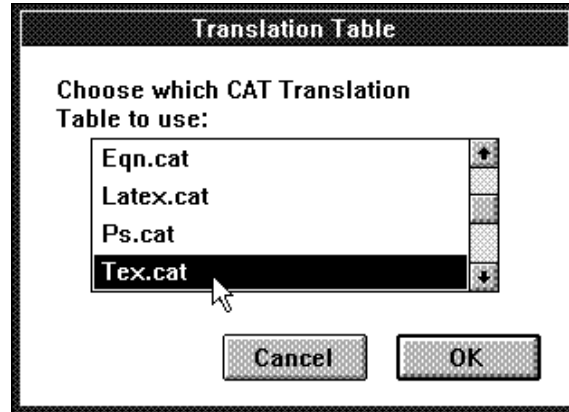
<u>Translation Table</u>	<u>Intended Use</u>
AMIPRO.CAT	For use with Lotus Ami Pro's proprietary equation markup language
WP.CAT	For use with WordPerfect's proprietary equation markup language
EQN.CAT	For use with eqn and troff systems
PS.CAT	Produces PostScript descriptions of expressions, complete with EPS headers
TEXT.CAT	For copying paragraphs of text and learning how the CAT system works
TEX.CAT	For use with T _E X systems
LATEX.CAT	For use with the L ^A T _E X macro package developed by Leslie Lamport
AMSTEX.CAT	For use with the AMST _E X macro package, a special macro set from the American Mathematical Society

If you modify a table, be sure to make a copy of the original CAT file.

QUICK CAT TOUR

Choosing a Translation Table

Choose **Translation Table...** from the **Options**→**Copying...** submenu. A dialog appears with a list of translation tables.



Click on a translation table (e.g., **TEX.CAT**) in the scrolling list, then click **OK**.

Setting for Text Copying

Choose **Copy As Text** from the **Options**→**Copying...** submenu. This tells *Expressionist* to copy text instead of pictures when a Cut or Copy command is issued.

Copying

Select the expression to be copied. (Drag or use Select All.)

Cut or Copy. *Expressionist* puts a text description of the selection in the clipboard.

Pasting

Go to an application (e.g., a text editor or a word processor) and Paste the clipboard. Examine the text exported by *Expressionist*.

```
%]]Expr[#>`b___))# b(<" *~: ;bP8&c55*V <c!$1^P} "!Symbol^:&c0 .Z|
%|: &c55*/0 <c" #(&&c55)E :!&c0 /7: &c55* d&c55)l)^;&&c552O^P}}# b D
b!( b!L!WW}}]]
```

```
$$V\ \left({P}\right)\ \rm \equiv \rm \mit \ -\
\int_{0}^{P}\bf E\ \rm \rm \cdot \rm \mit \ d\bf l$$
```

Your text will be different depending on your expression and the selected translation table. It may look like random nonsense now, but as you will see, it all makes sense.

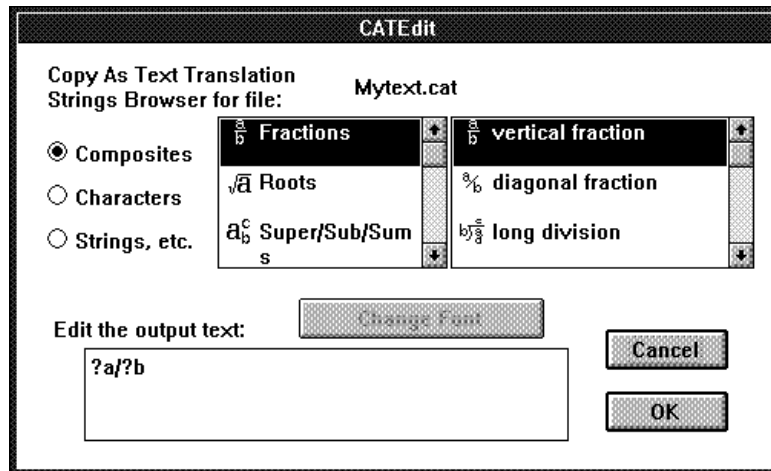
MODIFYING A TRANSLATION TABLE

You can modify a translation table so that it outputs a different text description for one or more composite, string, or character. The following pages take you through the various areas of a CAT translation table, making simple modifications as you go.

Instead of modifying the TEXT.CAT file directly, make a copy of it so you will be modifying the copy. Go to the DOS prompt and enter the command **CD C:\EXPR\EXPRPREF**. Then, enter the command **COPY TEXT.CAT MYTEXT.CAT**. When the copy is completed, return to Windows and start *Expressionist*.

From *Expressionist*, choose **MYTEXT.CAT** as the current Translation Table. Next, make sure *Expressionist* is set to Copy as Text. (You learned how to do these steps in the Quick CAT Tour.)

Expressionist's Translation Browser is used to view and/or modify the current CAT output descriptions. Choose **Translation Browser...** from the **Options** → **Copying...** submenu.



This is a browser which shows you the output text that *Expressionist* generates for each possible equation element. Above, it is showing the output text for a fraction, “?a/?b”. The “?a” represents the numerator and the “?b” is for the denominator. The text output for $\frac{1}{2}$ would be “1/2”.

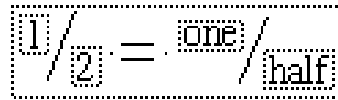
Notice the three radio buttons on the left side. The **Composites** button lets you browse the typical *Expressionist* structures, including fractions, fences, integrals, and the like. The **Characters** button lets you browse individual characters, including pseudocharacters and font changes. The **Strings, etc.** button lets you browse other output text.

The scrolling list to the right of the radio buttons lets you choose a more specific category, such as fence composites or pseudocharacters. The scrolling list on the far right is for choosing a specific kind of element in the category, such as a long division composite or a script ell pseudocharacter.

The box at the bottom of the Translation Browser is for typing in the sequence of characters to be output for the item selected in the right scroll box.

Modifying Composites

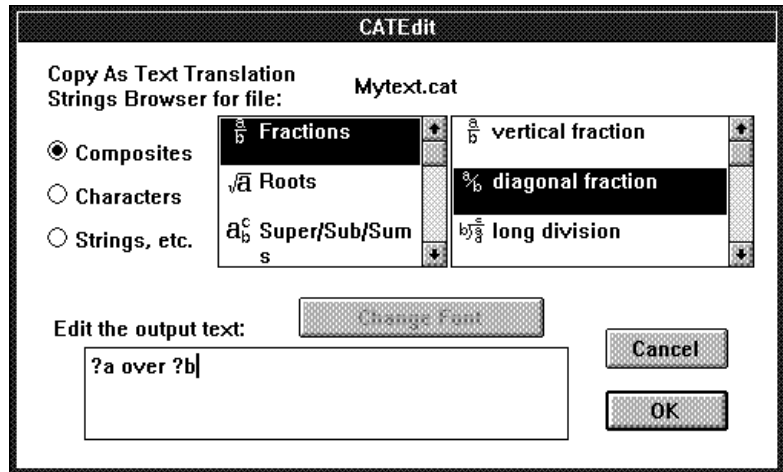
In this example we will change the output text for a diagonal fraction composite. Enter a diagonal fraction in an expression window:



Copy the fraction out of *Expressionist*, then Paste into your editor.

$$1/2 = \text{one/half}$$

Go back to the Translation Browser and choose **Composites, Fractions, diagonal fraction**. Change the output string to **?a over ?b**.



Click **OK**, then select and copy out the same expression. Go to your editor and paste to get:

$$1 \text{ over } 2 = \text{one over half}$$

Note that in the Translation Browser's output string the numerator was indicated by ?a and the denominator was ?b.

In this example, we examine the behavior of a composite.

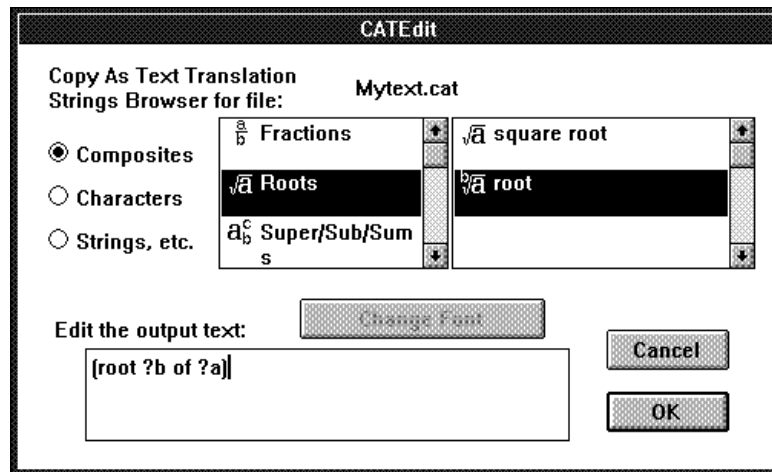
Make the expression:



Copy it and Paste it into your text editor. You get:

$$(8)^{(1 / (3))}$$

Use the Browser to select **Composites**, **Roots**, **root**, and change the output string to **(root ?b of ?a)**.



Click **OK**, then select and copy out the same expression. Go to your editor and paste to get:

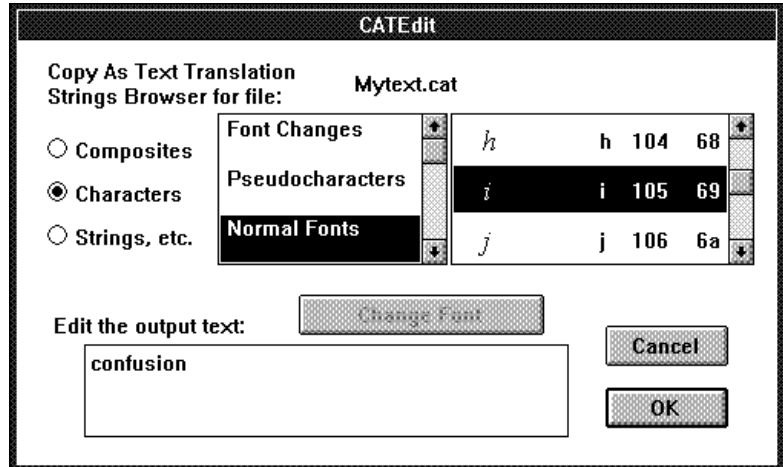
$$(\text{root } 3 \text{ of } 8)$$

Next, go back to the expression window and delete the 3 from the expression. Select and copy. Go to your editor and paste to get:

$$\text{sqrt}(8)$$

Because the exponent string of the root composite was empty, *Expressionist* used a different item from the translation table. Something similar occurs with subscripts and superscripts.

In this example, we are going to redefine the character “i”. Select **Characters**, **Normal Fonts**, *i* from the Translation Browser. Edit the output string, replacing **i** with the word **confusion**.



Click **OK**, then create this expression:

`quick fox`

Select it, copy it, and paste it into your editor.

`quconfusionck fox`

The “i” in the word “quick” was replaced with the new output string for “i” you put in the Browser. This feature can lead to unexpected results, so use with caution.

A more likely and useful implementation of this technique is to modify the definition of a special symbol. For instance, you could change the • character in Symbol font to the equivalent T_EX macro `\infty`.

Symbol and Other
Fonts

If, instead of clicking on **Normal Fonts**, you click on **Symbol Font**, you will access the output text for all of the characters in the Symbol font. This is an entirely independent set which is used only when Symbol font characters are used in an expression.

If you have other special fonts, you can have *Expressionist* recognize them for special treatment also. With the **Characters** radio button set, click on a blank entry below **Symbol**, click the **Change Font** button, and choose the special font.

If the font or style is changed inside an expression, these font changes may be defined in the translation table.

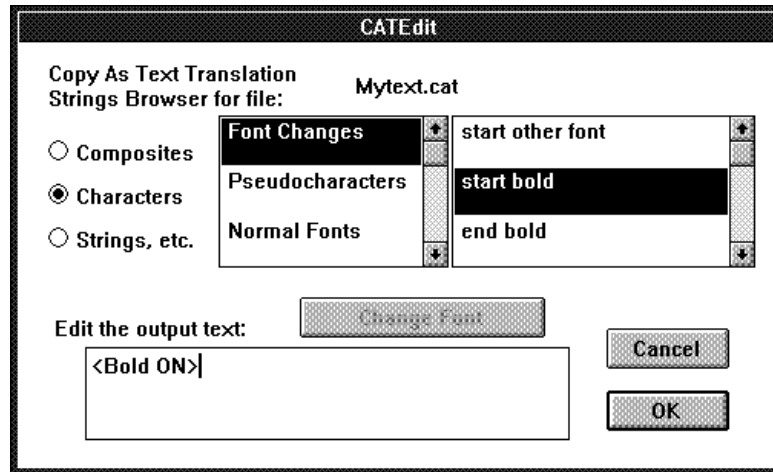
Create the following expression (with the styles applied as shown):

brown fox jumps over

Select it, copy it, and paste it into your editor. The resulting text does not have any font styles (which is the nature of plain text):

brown fox jumps over

Go into the browser and select **Characters, Font Changes, start bold**. Edit the output text—which is nonexistent for this item—to **<Bold ON>**.



Edit the **end bold** item to **<Bold OFF>** and do likewise for the italic and underline font changes. Click OK, then select and copy out the same expression. Go to your editor and paste to get:

<Italic OFF>brown <Bold ON>fox<Bold OFF> <Italic ON>jumps<Italic OFF> <Underline ON>over

(The default font can affect the style changes. For instance, this was made with italic default, it was all changed to plain, then styles were applied individually.)

There are many other kinds items that you can tinker with, most of which work like the examples in this chapter. There are some exceptions and some special circumstances that you should be aware of. You might have to experiment with these to see how they all work for your situation.

Many composites have alternative special cases. For instance, integrals have many variations. All marks are listed twice; once for above and once for below.

Fences, because they have two independent sides that can be mixed and matched, each invoke two concentric entries. If the output text for “parenthesis on left” were “((?a”, and the text for “bracket on right” were “?a]]”, the expression (5] would copy as “((5]]]”.

Matrices, tensors, and trees work differently because they are variable-sized. Their output text comes in three or four pieces. For matrices there is text to start a matrix (perhaps an opening bracket), text to separate columns (perhaps a comma or tab), text to separate rows (perhaps a semicolon or return), and text to end (perhaps a closing bracket). For tensors and trees there is text to start, which takes the base as an argument, then there is text to be repeated once for each index or child, then text to end.

To make a special case for a word that is a function name, such as “log”, go to **Characters, Font Changes**, and choose **Function Name**. The argument is the name itself. Normally this is ?a, which has no effect. T_EX, on the other hand, uses “\?a” which makes it so that “sin” comes out as “\sin”, which is a T_EX convention.

Be careful when using this powerful item.

The most powerful output text item is under **Strings, Global, Layout**. This defines the entire expression string that you copy out. Normally this is “?a”, to just pass through what you intended. If you wanted to enclose all equations in dollar signs, you would change this to “\$?a\$”. If you changed the output text to say “oxford”, with no ?a or anything, you would always get the result “oxford” regardless of your expression. The *Expressionist* equation encoding can be included by adding ?b. If this part of the copied text is pasted back into *Expressionist*, your equation will be recreated. (Hiding this encoding in your destination text—usually by commenting it out—is an issue you will have to deal with). The code ?c refers to the PostScript rendition of your equation, as if you had generated an EPS text file.

The encoding line break can be used to break up long lines in case your target system cannot deal with long lines of text.

Special Symbols

The *Expressionist* Copy As Text facility uses special codes to represent various elements of expressions. You have probably already figured out that a question mark followed by the letter “a” is used for the first argument. The codes used are as follows:

All codes start with a question mark. To output a question mark, enter two together.

<u>Code</u>	<u>Purpose</u>
?.	new line
?,	tab
{?	start font grouping
?}	end font grouping
?a, ?b, ?c, ?d	arguments of <i>Expressionist</i> macros/expressions

Working with T_EX

INTRODUCTION

TEX is a special document typesetting language renowned for programmability and its high quality output. TEX was designed by Donald Knuth of Stanford University during the late 1970s and early 1980s. It was designed for the highest typesetting quality, high functionality and programmability, consistency, reliability, and ease of use relative to contemporary systems.

Although TEX is not WYSIWYG (what you see is what you get), TEX has a large following, especially since documents prepared on one computer system can be easily moved to other computer systems running TEX. It is very popular in academic environments where mathematical typesetting is required, and where the structure of documents is complex enough to require programmability using macros.

Using *Expressionist*, you can create equations in a WYSIWYG manner for use with TEX programs. This section describes how to do so. If you installed *Expressionist* as described in the installation chapter and you read the section on Copying Text, you are ready to go.

Three TEX translation tables are supplied with *Expressionist*. These CAT files in the EXPRPREF directory are: AMSTEX.CAT, LATEX.CAT, and TEX.CAT. The appropriate translation table should be used for the TEX you are using, i.e., TEX.CAT for standard TEX expressions, AMSTEX.CAT and LATEX.CAT for the AMSTEX and L^ATEX macro packages respectively.

Each TEX translation table exports equations in display style. If you wish to change them to inline style, instructions to do so follow. The text generated by *Expressionist* includes a comment that describes the encoding of the mathematical expression. This can be removed as explained in a later example.

For more information on TEX, contact: TEX Users Group, P.O. Box 869, Santa Barbara, California 93102. Phone (805) 963-1338, fax (805) 963-8358, or e-mail tug@math.ams.org.

Prescience recommends the following references:

- Donald Knuth, *The TEXbook*, Addison-Wesley, 1986
- Leslie Lamport, *L^ATEX*, Addison-Wesley, 1986
- Michael Spivak, *The Joy of TEX*, AMS, 1986
- Jane Hahn, *L^ATEX for Everyone*, Prentice-Hall, 1993

TEX.CAT Notes

Equations are set using the standard $\$$ to begin and end mathematical expressions. The translator does not know when to use $\backslash\text{big}$, $\backslash\text{bigg}$, etc. for creating large delimiters, e.g., parentheses, braces, etc.

AMSTEX.CAT Notes

Equations are set using the standard $\$$ to begin and end mathematical expressions in display style. You must modify arrays to generate the correct alignment for items. (See the example in the following L^AT_EX notes.)

LATEX.CAT Notes

Equations are set using the standard $\left[$ to begin and $\right]$ to end mathematical expressions. You must modify arrays to generate the correct alignment for items inside the array. The following example explains the necessary changes.

$$\left(\begin{array}{cc} \cos(\theta) & 1 \\ 1 & \sin(\theta) \end{array} \right)$$

```
 $\left(\begin{array}{2c}\cos(\theta \text{ \rm })&1\\1&\sin(\theta \text{ \rm })\end{array}\right)$ 
```

The $\{2c\}$ following $\{array\}$ must be changed to describe the preferred column alignment. For example, if you want right-aligned columns, change the $2c$ to rr . Currently, the T_EX translator emits the number of columns and a c to specify the columnar alignment.

```
 $\left(\begin{array}{rr}\cos(\theta \text{ \rm })&1\\1&\sin(\theta \text{ \rm })\end{array}\right)$ 
```

Function Macros

Special T_EX macros for mathematical functions are included in the file `FMACROS.TEX`. If you want to use some of the non-standard functions that *Expressionist* knows about, but T_EX does not, copy this file into your T_EX inputs directory. These are special T_EX macros for symbols and functions that are not defined in standard versions of T_EX. Examples include " and arc (for use in functions like "arc tan").

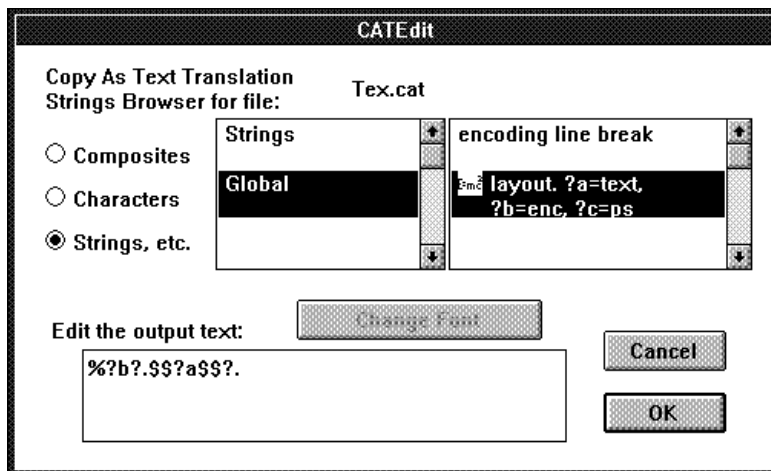
Changing to Inline Equation Style

To change from display style to in-line style there are two choices:

- Edit the T_EX code generated by *Expressionist* every time you need an inline equation.
- Modify the T_EX translation table once, then modify the T_EX any time you need equations set in display style.

To modify the T_EX translation table so that in-line expressions are generated instead of display equations, choose **Translation Browser...** from the **Options** \blacktriangleright **Copying...** submenu. In the Translation Browser

dialog, click the **Strings, etc.** radio button, then choose the **Global** category, then choose \emc layout from the right scroll box.



Change the output string from

`%?b?.$\$?a\$$?.`

to

`%?b?.$?a$?.`

After making this change to your \TeX translation table, all copied expressions are ready for use as in-line equations.

Encoding

When text is copied out of *Expressionist*, the default translation tables generate a \TeX comment and \TeX code that describes the expression. The comment includes encoding necessary to paste the equation back into *Expressionist*.

Consider the following simple expression:


$$\int x dx$$

The \TeX code generated by *Expressionist* for this expression is:

```
%]|Expr|[#>`b____)!# b$.<c" #(#" *~: ;bP8&c552x
dx}__}]# b D b!( b!L!WW}]|[
$\int_{ }^{ }\mit x\ dx$
```

The percent sign is the comment delimiter. The sequence of characters starting with `]|Expr|` and ending with `]|` is the special *Expressionist* encoding. The remainder is \TeX code.

Removing the Encoding

If you do not need to move TEX expressions back into *Expressionist* and do not want the *Expressionist* comment cluttering up your TEX document, you can remove the encoding by editing the translation table. Choose **Translation Table...** from the **Options** **Copying...** submenu. In the Translation Browser dialog, click the **Strings, etc.** radio button, then choose the **Global** category, then choose  layout from the right scroll box.

Edit the output string, changing it from

$$%?b?. $$?a$$? .$$

to

$$$$?a$$? .$$

for display equations. Or, for inline equations, change it from

$$%?b?. $?a$? .$$

to

$$$?a$? .$$

When the expression is copied out of *Expressionist*, the following text is in the clipboard:

$$\int_{ }^{ } x \, dx$$

Clearly, this is much shorter than the code originally generated. Just remember you can not move this equation back to *Expressionist* to modify your equations if you make this change.

Translation Limitations

The TEX translator can not capture all the control provided by TEX. The following example illustrates this.

Consider the following equation,

$$n/\log n$$

and the TEX code generated by *Expressionist*,

$$$$n/\log n$$$$

Use *Expressionist*'s tweaking features to move various characters around as shown below.

$$n / \log n$$

The TEX code generated is the same as shown previously because *Expressionist*'s TEX translators do not translate any tweaking within an equation.

$$n^{\log n}$$

If you are using *Expressionist* to tweak equations, please note that the TEX translators do not include this extra descriptive information in the resulting TEX code.

Acknowledgments

We would like to thank Diane Heining and Lance Carnes of Personal TEX, Inc. for supplying us with the most recent version of PC TEX and for their assistance in testing *Expressionist*. For more information regarding the PC TEX software, contact:

Personal TEX, Inc.
12 Madrona Street
Mill Valley, CA 94941
phone (415) 388-8853
fax (415) 388-8865