https://www.quora.com/What-is-the-difference-between-Markov-Random-Fields-MRFs-and-Conditional-Random-Fields-CRFs-When-should-I-use-one-over-the-other  
Just a follow-up on Eren's answer.  
  
Let X denote a multi-dimensional input (i.e., the features), and let Y denote a multi-dimensional output (i.e., the structured label).  Typically, we have the features x, and are interested in predicting the distribution of the label P(Y|X=x), or the best possible label argmax\_y P(Y=y|X=x).    
  
A CRF is essentially a structured extension of logistic regression, and models the CONDITIONAL probabilities P(Y|X).  It does not model anything else.  
  
A MRF models the JOINT probabilities of both Y and X simultaneously.  In other words, it models P(Y,X).  You can do various tricks to compute P(Y|X=x) for a given input x.    
  
The advantage of CRFs is that they are more directly modeling the standard prediction problem P(Y|X=x).  As such, they're often more accurate.  But they cannot (without further modifications) be used for anything other than the standard prediction problem.  
  
The advantage of MRFs is that they are fully generative, and so can model arbitrary prediction problems.  For instance, suppose for some reason that you had missing values in your input x.  Then an MRF can marginalize over the missing values because it models the full probability distribution.  As another example, suppose you had the y and wanted to predict X.  Then an MRF can also predict P(X|Y=y), but a CRF cannot.  
  
For me, CRFs are generally much more useful than MRFs, because I'm generally interested in the standard prediction problem P(Y|X=x).

Different people use the terms in different ways.  The following helped me, at least, understand:  
  
A CRF \*is\* an MRF, where the factor potentials are defined by some conditioned-on data.  
  
An MRF is just a probability distribution over random variables.  There are no inputs or outputs.  Sometimes when people say MRF, they're talking about one single network.  Sometimes, they mean a joint distribution over both inputs and outputs.  I don't think the terminology is totally consistent (at least in the context of machine learning).  
  
CRF's always involve inputs and outputs.  Typically, a CRF is used for structured prediction where you train it on many (input,output) pairs.  Given an input, the CRF's parameters define a probability distribution over possible outputs.  So in the data, every pair has its own little MRF; what the MRF's have in common are shared parameter weights.  
  
Another name for CRF might be "a generalization of logistic regression to have structured outputs, like chains, trees, or grids."  
  
That's why sometimes it's written, an MRF is just p(y), while a CRF is p(y | x).  Personally, I think it's not totally accurate to say a CRF conditions on the input (it's not really a random variable, it's more like, data that defines the factor potentials...), but that notation can be helpful in getting across what's going on.  It's certainly the case that a CRF does not define a probability distribution over inputs.

Conditional Random fields are directly formulate to compute p(label | instance) instead of modelling p(label, instances) as a generative model. In general each node on the CRF graph is computed via a disciminative model like logistic regression and at the upper layer correlations of those nodes are considered.  
  
You can follow that schema at Markov Random Field but then it is called CRF :).

------------------------------------------------------------------------------------------------------------I am adding a small section on discriminative vs generative. Generative models are based on model of joint distribution [math]p(*y*,**x**)[/math] while discriminative is based on [math]p(*y*|**x**)[/math]. In generative models you model argmaxyp(x|y)p(x)argmaxyp(x|y)p(x) while in discriminative models you model just  
argmaxyp(y|x)argmaxyp(y|x).

-------------------------------------------------------------------------------------------------------

<https://stats.stackexchange.com/questions/156697/whats-the-difference-between-a-markov-random-field-and-a-conditional-random-fie>

Conditinal Random Fields (CRFs) are a special case of Markov Random Fields (MRFs).

1.5.4 Conditional Random Field

A Conditional Random Field (CRF) is a form of MRF that defines a posterior for variables x given data z, as with the hidden MRF above. Unlike the hidden MRF, however, the factorization into the data distribution P (x|z) and the prior P (x) is not made explicit [288]. This allows complex dependencies of x on z to be written directly in the posterior distribution, without the factorization being made explicit. (Given P (x|z), such factorizations always exist, however—infinitely many of them, in fact—so there is no suggestion that the CRF is more general than the hidden MRF, only that it may be more convenient to deal with.)

Source: Blake, Kohli and Rother: Markov random fields for vision and image processing. 2011.

A conditional random field or CRF (Lafferty et al. 2001), sometimes a discriminative random field (Kumar and Hebert 2003), is just a version of an MRF where all the clique potentials are conditioned on input features: [...]

The advantage of a CRF over an MRF is analogous to the advantage of a discriminative classifier over a generative classifier (see Section 8.6), namely, we don’t need to “waste resources” modeling things that we always observe. [...]

The disadvantage of CRFs over MRFs is that they require labeled training data, and they are slower to train[...]

Source: Kevin P. Murphy: [Machine Learning: A Probabilistic Perspective](https://www.cs.ubc.ca/~murphyk/MLbook/pml-print3-ch19.pdf)

Answering my question:

If I fix the values of the observed nodes of an MRF, does it become a CRF?

Yes. Fixing the values is the same as conditioning on them. However, you should note that there are differences in training, too.

Watching many of the [lectures about PGM](https://www.coursera.org/course/pgm) (probabilistic graphical models) on coursera helped me a lot.

<http://www.cs.sfu.ca/~oschulte/teaching/726/spring11/slides/mychapter4.pdf>

https://stats.stackexchange.com/questions/58221/intutive-difference-between-hidden-markov-models-and-conditional-random-fields